

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 08:38:45 ; Search time 12.9816 Seconds
(without alignments)
1851.363 Million cell updates/sec

Title: US-09-730-465-2
Perfect score: 1326
Sequence: 1 MWAGSDAGRALGVLSVWCLL.....VLVMGILKCDRKRPRTNSN 250
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1326	100.0	250	2 A28564	lymphocyte functio
2	131.5	9.9	240	2 JL0143	antigen BCM1 precu
3	126	9.5	347	2 S41638	T-cell surface gly
4	123	9.3	243	2 A53244	leukocyte antigen
5	115.5	8.7	351	1 RWHUC2	T-cell surface gly
6	106	8.0	490	2 T38088	protoporphyrinogen
7	105.5	8.0	773	1 QRRBG	secretory componen
8	103	7.8	240	2 S01299	OX-45 membrane gly
9	100.5	7.6	1434	2 C90109	DNA-directed RNA p
10	97.5	7.4	276	2 A71653	hypothetical prote
11	95	7.2	344	2 RWRFC2	T-cell surface gly
12	95	7.2	2489	2 S59782	probable membrane
13	93.5	7.1	740	2 B84741	hypothetical prote
14	93	7.0	902	2 S61144	glycogen phosphory
15	92.5	7.0	299	2 I46690	CD80 precursor - r
16	91.5	6.9	1437	2 T31093	probable protein-t
17	91	6.9	289	2 G00031	B7 protein - red-c
18	91	6.9	629	2 A46500	Ly-9.2 antigen - m
19	90.5	6.8	276	2 S20690	31.6k hypothetical
20	90.5	6.8	1113	2 T14260	period protein Per
21	89.5	6.7	546	2 T26568	hypothetical prote
22	89	6.7	873	1 I48952	VLDL receptor prec
23	89	6.7	880	2 T21538	hypothetical prote
24	88.5	6.7	309	2 I49503	B-lymphocyte activ
25	88.5	6.7	750	1 D69403	conserved hypothet
26	88	6.6	690	2 S54775	cell size regulati
27	87.5	6.6	272	2 I48268	biliary glycoprote
28	87.5	6.6	293	2 T49947	hypothetical prote
29	87	6.6	288	2 A45803	B-cell-restricted

30	87	6.6	1453	2 P88640	protein F52C12.4 [
31	86.5	6.5	398	2 I49443	gene 2B4 protein -
32	86	6.5	836	2 E97851	hypothetical prote
33	86	6.5	1018	2 JC4211	neural adhesion pr
34	86	6.5	1403	2 S64142	hypothetical prote
35	85.5	6.4	216	2 S29258	Ig lambda chain V
36	85.5	6.4	526	2 S46766	hypothetical prote
37	85.5	6.4	767	1 WMBEP6	ribonucleoside-dip
38	85.5	6.4	793	1 JC6161	kinesin-associated
39	85	6.4	763	2 T49089	hypothetical prote
40	85	6.4	785	1 QOV28	D5 protein - vacci
41	85	6.4	785	2 E72162	FSR protein - vari
42	85	6.4	785	2 D42515	DSR protein - vacc
43	85	6.4	785	2 B36847	ATP/GTP-binding pr
44	85	6.4	785	2 T28533	hypothetical prote
45	85	6.4	785	2 T37378	probable 90.4K pro

ALIGNMENTS

RESULT 1

A28564

lymphocyte function-associated antigen 3, transmembrane splice form precursor - human

N:Alternate names: CD58 antigen; surface glycoprotein LFA-3

C:Species: Homo sapiens (man)

C>Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 02-Aug-2002

C:Accession: A28564; S01269

C:Wallner, B.F.; Frey, A.; Tizard, R.; Mattaliano, R.J.; Hession, C.; Sanders, M.E.

J. Exp. Med. 166, 923-932, 1987

A:Title: Primary structure of lymphocyte function-associated antigen 3 (LFA-3). The 1

A:Reference number: A28564; MUID:88009714; PMID:3309127

A:Accession: A28564

A:Molecule type: mRNA

A:Residues: 1-250 <NAL>

A:Cross-references: GB:Y00636; NID:g34346; PIDN:CAA68668.1; PID:g34347

A:Experimental source: erythrocytes

A:Note: the authors translated the codon TAT for residue 34 as Val

R:Seed, B.

Nature 329, 840-842, 1987

A:Title: An LFA-3 cDNA encodes a phospholipid-linked membrane protein homologous to 1

A:Reference number: S01269; MUID:88039074; PMID:3313052

A:Accession: S01269

A:Molecule type: mRNA

A:Residues: 1-235, 'VL' <SBE>

A:Cross-references: EMBL:X06296; NID:g34349; PIDN:CAA29622.1; PID:g34350

C:Comment: For an alternative splice form, see PIR:S01269.

C:Comment: For an alternative splice form, see PIR:A28564.

C:Genetics:

A:Gene: GDB:CD58; LFA3

A:Cross-references: GDB:120580; OMIM:153420

A:Map position: lp13-lp13

C:Superfamily: human B-cell antigen CD19; immunoglobulin homology

C:Keywords: alternative splicing; glycoprotein; phosphatidylinositol linkage; surface

F:1-30/Domain: signal sequence #status predicted <SIG>

F:31-250/Product: lymphocyte function-associated antigen 3 transmembrane splice form

F:135-189/Domain: immunoglobulin homology <IMM>

F:40.94,109,135,169,195/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 1326; DB 2; Length 250;

Best Local Similarity 100.0%; Pred No 2.3e-105;

Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWAGSDAGRALGVLSVWCLLHCFGISCFSSQIYGVVGNVTFHVSNNVPLKEVLWKKQK 60

Db 1 MWAGSDAGRALGVLSVWCLLHCFGISCFSSQIYGVVGNVTFHVSNNVPLKEVLWKKQK 60

QY 61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTYNLTSSDEDEYEMESPNTITDTMFFLYV 120

Db 61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTYNLTSSDEDEYEMESPNTITDTMFFLYV 120

QY 121 LESLSPPTLTCAITNGSIEVOQMGIPEHYNSHRGLIMYSWDCPMEOCKRNSTIYFKMEND 180

```
Db 121 LESLSPPTLTCALTNLSIEVQCMIPBHYNSHRGLIMYSWDCPMEQCKRNSTSTIYFKMEND 180
Qy 181 LPQKTOCTLSNPLNTSSIIITTCIPSSGHSRRHYALIPPLAVITTCIVLYMNGILKC 240
Db 181 LPQKTOCTLSNPLNTSSIIITTCIPSSGHSRRHYALIPPLAVITTCIVLYMNGILKC 240
Qy 241 DRKPDRTNSN 250
Db 241 DRKPDRTNSN 250

RESULT 2
JL0143
N:Antigen BCM1 precursor - mouse
N:Alternate names: CD48 antigen homolog sgp-60; OX45 antigen, Blast-1 antigen
C:Species: Mus musculus (house mouse)
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 31-Jan-2000
C:Accession: JL0143; S21319; A47469; B47469
R:Wong, Y.W.; Williams, A.F.; Kingmore, S.F.; Seldin, M.F.
J. Exp. Med. 171, 2115-2130, 1990
A:Title: Structure, expression, and genetic linkage of the mouse BCM1 (OX45 or Blast-1) region on mouse chromosome 3.
A:Reference number: JL0143; MUID:90278362; PMID:1693656
A:Accession: JL0143
A:Molecule type: mRNA
A:Residues: 1-240 <WON>
A:Cross-references: EMBL:X17501; NID:g50134; PIDN:CAA35542.1; PID:g50135
R:Wong, Y.W.; Williams, A.F.; Kingmore, S.F.; Seldin, M.F.
submitted to the EMBL Data Library, June 1990
A:Description: Structure, expression and genetic linkage of the mouse BCM1 (OX45 or Blast-1) region on mouse chromosome 3.
A:Reference number: S21319
A:Accession: S21319
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-240 <WON2>
A:Cross-references: EMBL:X53526; NID:g50138; PIDN:CAA37604.1; PID:g50139
R:Cabrero, J.G.; Freeman, G.J.; Lane, W.S.; Reiser, H.
Proc. Natl. Acad. Sci. U.S.A. 90, 3418-3422, 1993
A:Title: Identification, by protein sequencing and gene transfection, of sgp-60 as the mouse BCM1 (OX45 or Blast-1) region on mouse chromosome 3.
A:Reference number: A47469; MUID:93234508; PMID:8475091
A:Accession: A47469
A:Molecule type: protein
A:Residues: 74-80 <CAB>
A:Experimental source: EL-4 lymphoma cells
A:Note: sequence extracted from NCBI backbone (NCBIP:129658)
A:Accession: B47469
A:Molecule type: protein
A:Residues: 84-98 <CA2>
A:Experimental source: EL-4 lymphoma cells
A:Note: sequence extracted from NCBI backbone (NCBIP:129660)
C:Comment: This antigen is widely expressed on leukocytes and is likely to be anchored to the cell surface.
C:Superfamily: B-cell surface glycoprotein blast-1
C:Keywords: blocked carboxyl end; glycoprotein; lipoprotein; phosphatidylinositol linkage
F:1-22/domain: signal sequence #status predicted <SIG>
F:23-217/product: antigen BCM1 #status predicted <MAT>
F:218-240/domain: carboxyl-terminal propeptide #status predicted <CPT>
F:32,38,70,136,203/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:217/Modified site: GPI-anchor ethanolamine amidated carboxyl end (Ser) (in mature form)

Query Match 9.9%; Score 131.5; DB 2; Length 240;
Best Local Similarity 21.9%; Pred. No. 0.00087;
Matches 57; Conservative 36; Mismatches 90; Indels 77; Gaps 11;
Qy 13 VLSVVCVLLHCFGIFSCFSQIYGVYGNVTFHVPSN--PLKEVLMKKQKQ-KVAELE-N 68
Db 11 VLELLPLGTFQGHSGIPDINATTCGNSVTLKIHKDPGLPGYKRTITWLHPTKNGKILEYNN 70
Qy 69 SEERAF--SSFKNRVYLDTVSGSLTIYNTLTSDEDEYEMESPNTIDP--MKFFLYLVLESIPS 126
Db 71 STKIFSEFGKRVYLEENNGALHNSVRKDKGTYYRVRLRETENELKITLEVDPVPVK 130
Qy 127 PTLTCALTNLS-----IEVQCMIPBHYNSHRGLIMYSWDCPMEQCKRNSTSTIYFKMENDL 181
```

```
Db 131 PSIEINKTEASTDSCHLRSLCEVQDQ-----VDYTW-----YESSGPF 169
Qy 182 PQK-----IOCTLSNPLNTSSIIITTCIPSSGHSRRHYALIPPL 223
Db 170 PKKSPGVLDLIVTPQNKSTFYTCQVSNPVSSKNDTV-----YFTLPCDL 214
Qy 224 A-----VITTCIV 231
Db 215 ARSSGVCTATWLVLTLLII 234

RESULT 3
S41638
N:cell surface glycoprotein CD2 precursor - horse
N:Alternate names: T-lymphocyte surface antigen CD2
C:Species: Equus caballus (domestic horse)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 23-Jul-1999
C:Accession: S41638; S31578
R:Tavernier, A.S.; Kydd, J.H.; Bodian, D.L.; Jones, E.Y.; Stuart, D.I.; Davis, S.J.; B. Eur. J. Biochem. 219, 969-976, 1994
A:Title: Expression cloning of an equine T-lymphocyte glycoprotein CD2 cDNA. Structure and function of the protein.
A:Reference number: S41638; MUID:94155904; PMID:7906650
A:Accession: S41638
A:Molecule type: mRNA
A:Residues: 1-347 <TAV>
A:Cross-references: EMBL:X69884; NID:g1057; PIDN:CAA49511.1; PID:g1058
C:Superfamily: T-cell surface glycoprotein CD2
C:Keywords: glycoprotein; surface antigen; T-cell; transmembrane protein
F:1-24/domain: signal sequence #status predicted <SIG>
F:25-347/product: T-cell surface glycoprotein CD2 #status predicted <MAT>

Query Match 9.5%; Score 126; DB 2; Length 347;
Best Local Similarity 23.3%; Pred. No. 0.0041;
Matches 51; Conservative 44; Mismatches 72; Indels 52; Gaps 11;
Qy 14 LSVVC-LLHCFGIFSCFSQ-----YGVYGNVTFHVPS---NPLKEVLMKKQK 60
Db 1 MNACKLLASFLILIFFSSKGA VKNNITLIGALERDINLDIPAFQMSERHVEDIQWSKKG 60
Qy 61 DKVAELENSSEAFSSFKNRVYLDTVSGSLTIYNTLTSDE-----DEVEMESPNTIDTK 115
Db 61 TKIAKEFG--SMTFQDKTYEVLKNGTLKIKHLERIHGTYKYVDAYDSGKNVLEE-T 116
Qy 116 FFYLVLESPLSP-----TLTCALTNLS-IEVOCMIPDH--YNSHRGLIMYSWDCPM 163
Db 117 FHLSLLEMYSKPNISWSCTNTTLTCEYTKGTDFELKLYLNGRMQKSPRKVIYV----- 170
Qy 164 EQCKRNSTSTIYFKMENDLPQKIOCTLSNPLNTSSIIIL 202
Db 171 ---KRASNOI-----ASFKCTANNTVSESSSVVI 197

RESULT 4
A53244
N:leukocyte antigen CD48 precursor - human
N:Alternate names: B-cell surface glycoprotein Blast-1; CD48; leukocyte antigen HuLy-4
C:Species: Homo sapiens (man)
C>Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 23-Jul-1999
C:Accession: A53244; A53245; A27236; A39689; A54320
R:Korinek, V.; Stefanova, I.; Angelisova, P.; Hilgert, I.; Horejsl, V. Immunogenetics 33, 108-112, 1991
A:Title: The human leukocyte antigen CD48 (MEM-102) is closely related to the activated T-cell antigen CD44.
A:Reference number: A53244; MUID:91153857; PMID:1999350
A:Accession: A53244
A:Molecule type: mRNA
A:Residues: 1-243 <KOR>
A:Cross-references: GR:M37766; NID:g187518; PIDN:AAA36211.1; PID:g187519
A:Note: part of this sequence, including the amino end of the mature protein, was determined by sequencing of cDNA clones for CD48.
R:Vaughan, H.A.; Henning, M.W.; Purcell, D.F.J.; Mckenzie, I.F.C.; Sandrin, M.S. Immunogenetics 33, 113-117, 1991
A:Title: The isolation of cDNA clones for CD48.
A:Reference number: A53245; MUID:91153858; PMID:1999351
```


Db 143 TLTCEVMNGP-----DPELNLYQDGKHLKLSQRVITHKW-----TTS----- 179

Qy 180 DLPKIOCTLSNPLENTTSSILTTCTPSSGHSRHRVLTPI-----PLAVITTCIVLYM 234

Db 180 -LSAKRCKTAGNKV-SKESSVEPVSC-PKSG--LDYLIIGICGGSLMVFVALLVFI 234

Qy	235	NGILKCDRKPDRTNSN	250
Db	235	T-----KRKKQSRSRN	245

RESULT 6

T38088

protoporphyrinogen oxidase - fission yeast (*Schizosaccharomyces pombe*)

C.Species: Schizosaccharomyces pombe
C.Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C.Accession: T38088
R.Gentiles, S.; Odell, C.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.;
submitted to the EMBL Data Library, November 1995
A.Reference number: 221768

A;ACCESSION: T38088
A:STATUS: preliminary: translated from CB/EVAL/DBRT

RESULT 7

QRBG

secretory component precursor - rabbit

N:Alternate names: poly-Ig receptor; polymeric immunoglobulin receptor
N:Contents: free secretory component; transmembrane secretory component
C:Species: *Oryctolagus cuniculus* (domestic rabbit)
C:Date: 15-Nov-1984 #sequence_revision 15-Nov-1984 #text_change 16-Jul-1999
C:Accession: A02111; A28077
R:Mostov, K.E.; Friedlander, M.; Blobel, G.

R:Frutiger, S.; Hughes, G.J.; Hanly, W.C.; Jaton, J.C.
J. Biol. Chem. 263, 8120-8125, 1988

A:Title: Rabbit secretory components of different allotypes vary in their carbohydrate
A:Reference number: A28077; MUID:88228032; PMID:313139

A:Accession: A28077

A:Molecule type: protein

A:Residues: 87-114;410-424 <FRU>

C:Comment: This receptor binds polymeric IGA and IGM at the basolateral surface of ep
process, cleavage occurs to separate the extracellular portion, also known as the sec
C:Comment: The five domains exhibit homology with immunoglobulin V regions. The siml
C:Comment: Alternative splicing in the extracellular domain leads to high or low mole
C:Superfamily: secretory component; immunoglobulin homology

C:Keywords: alternative splicing; duplication; glycoprotein; immunoglobulin receptor;
F:18/Domain: signal sequence #status predicted <SIG>

F:19-773/Product: transmembrane secretory component #status predicted <NATM>

F:19-575/Product: free secretory component #status predicted <NATF>

F:30-647/Domain: extracellular #status predicted <EXT>

F:39-117/Domain: immunoglobulin homology <IM1>

F:148-227/Domain: immunoglobulin homology <IM2>

F:353-326/Domain: immunoglobulin homology <IM3>

F:362-440/Domain: immunoglobulin homology <IM4>

F:471-540/Domain: immunoglobulin homology <IM5>

F:648-670/Domain: transmembrane #status predicted <TMM>

F:671-773/Domain: intracellular #status predicted <INT>

F:46-115,155-225,260-324,369-438,478-538/Disulfide bonds: #status predicted
F:108/Binding site: carbohydrate (Asn) (covalent) #status experimental

F:108/Binding site: carbohydrate (Asn) (covalent) #status experimental

```

Query Match      8.0%; Score 105.5; DB 1; Length 773;
Best Local Similarity 22.0%; Pred No. 0.63;
Matches 47; Conservative 39; Mismatches 49; Gaps 11;

QY 39 GNWTHVP-----SNVPLKEVLWKQKDKVAELENSEFRFSSFNRRVYLDTVSG-----S 89
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 363 GSVTIRCPYPKKSDSHLQYLWEGSQRHLLDWDSGVLVQKDYGRALFEEPGNGTFS 422
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 90 LTIYLTSSDDEY---EMESPNIIDTMKFLVYLSLSPSTLT--CALTNNGSTVEOCMI 144
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 423 VYLNQLTAEDEGFYWCVSODDESLLTSYK--LQIVDGEPSITDFTAVOGEVETCHF 480
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 145 P-EHYNSHRGILIMYSDCPMEQCKNSTSYFKMENDLPKIQCTLSNPLNTFTSSIIIT 203
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 481 PCKYFESSEK---YW-----CKWNHDGCG-----EDLPTKL--SSSGDLVKCNLNVLIT 522
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 204 T-----CIPSSGSHSRHRYALIPIL 223
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 523 LTLDVSDEDEGWYWCAGKDGEHEFEVAAVRVEL 556
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

RESULT 8

501299

OX-45 membrane glycoprotein precursor - rat

N:Alternate names: MRC OX-45 antigen
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 23-Jul-1999
C:Accession: S01299
R:Killeen, N.; Moessner, R.; Arvieux, J.; Willis, A.; Williams, A.F.
EMBL J. 7, 3087-3091, 1988
A:Title: The MRC OX-45 antigen of rat leukocytes and endothelium is in a subset of rat lymphocytes and is related to the human CD45 antigen.
A:Reference number: S01299; PMID:89030603; PMID:1811129

A: Accession: 301239
A: Molecule type: mRNA


```
Best Local Similarity 27.9%; Pred. No. 1.7;
Matches 50; Conservative 26; Mismatches 69; Indels 34; Gaps 13;

QY 35 GVYVGNVTHVPSNVV-----LKEVLKKKQKDKVALENSSEAFSSFKNRVYLDTVS 87
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 26 GTVWVGALGRLNLPINPQMTDDIDEVRGSGTLVAEKKR-KMKPF--LKSGAFELAN 82
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 88 GSLLTYNTSSDEDEYME--SPNITDTM--KFFLYVLESPLSP-----TLTCAIT 134
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 83 GDLKIKNLTRDSDGTNVTVSTNGTRILDKALDLRILEMVSKPMIYWECSNATLCEVL 142
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 135 NSGIEVQCM---PEHNSHRGLIM-YSWD---CPMEQCKR-NSTSIYFKME-NDLPQK 184
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 143 EGT-DVELKLYQKGLHRSQRQKTMSTQNTLRAPF-RCKAVNRVSOESMEVYVNCPEK 199
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

RESULT 12
S59782
probable membrane protein YPR117w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein P8283.1
C:Species: Saccharomyces cerevisiae
C:Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 19-Apr-2002
C:Accession: S59782
submitted to the EMBL Data Library, July 1995
A:Description: The sequence of S. cerevisiae cosmid 8283.
A:Reference number: S59764
A:Accession: S59782
A:Molecule type: DNA
A:Residues: 1-2489 <NEL>
A:Cross-references: EMBL:U32445; NID:g914969; PID:g914988; GSPDB:GN00016; MIPS:YPR117w
C:Genetics:
A:Gene: MIPS:YPR117w
A:Cross-references: SGD:S0006321
A:Map position: 168
C:Keywords: transmembrane protein
F:16-32/Domain: transmembrane #status predicted <TM>
F:349-365/Domain: transmembrane #status predicted <TM>

Query Match 7.2%; Score 95; DB 2; Length 2489;
Best Local Similarity 21.2%; Pred. No. 22;
Matches 52; Conservative 39; Mismatches 84; Indels 70; Gaps 14;

QY 26 ISCFSQ-----LYGVVYGVNTEHVPSNVPLKEVLW-----KKQKDKVALENS 69
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 2168 LQCFNQENPVYTELLAPFENSYN--KNAPWEIFWILGPSVGGISDQDLIVSLQPL 2225
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 70 EFRAFSSFKNRVYLD-TVSGSLTYNTSSDEDEYME--SPNITDTM--KFFLYVLESPL 125
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 2226 IFK-----MDHKTSEKLMNLYFPKIEQTSIEPNSELPVPRSSTSFF-----SS 2269
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 126 SPTLTCAITNGSIEVQCMPIEYNSHRGLIMYSWDCPMEQCKRN-----STSIY 174
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 2270 SPVLRHSLNSGLSV-----YDAKD---VDSWDLRSIQKEGKKHGDHKKLSASLF 2319
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 175 FKWENDLPQKIOCTLSNPLNTSSII-----LTTCPSSGHSRHRVYALIPVAVTTTCI 230
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 2320 VQPDYNINEMVK--RSGTFFNFKSIIIRKTLMSVCYKGS-HS-----LLTDVNNLVIRVPV 2372
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 231 VLYMN 235
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 2373 LKYNH 2377
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

RESULT 13
B84741
hypothetical protein At2g33080 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: B84741
R:Lin.X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koob, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
```

```
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: B84741
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-740 <STO>
A:Cross-references: GB:AE002093; NID:g2924783; PIDN:AAC04912.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g33080
A:Map position: 2

Query Match 7.1%; Score 93.5; DB 2; Length 740;
Best Local Similarity 24.1%; Pred. No. 6.2;
Matches 49; Conservative 21; Mismatches 64; Indels 69; Gaps 7;

QY 70 EFRAFSSFKNRVYLDTVSGSLTYNTSSDEDEY-----EMESPNITDTMK 115
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 259 DLNLFSSLSKLTLYDLDSGNSISPRSLRS---DLYIPLTLEKLLLEQCGIIEFPNLIKTLQ 315
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 116 FFLYV-----LESLSPTL-----TCALTNGSIEVQCMPIEH 147
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 316 KLEYIDMSNNRINKIPEWLWRLPRLSRMSLANNSFNFGESGTDVLVNSSMEILFM---H 372
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 148 YNSHRGLI-----MYSDCPMEQCKRNSTSIYFKMENDLPQKIOCTLSNP 192
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 373 SNIIQALPNLPLSIKAFSAGYNNFSGEIPLSICNRSSLAALSPLPYNNTGKIPQCLSNL 432
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 193 LF-----NTTSSIIITTCIPSS 209
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 433 TFVHLRKNNEGSIPTDLCAGDS 455
   | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

RESULT 14
S61144
glycogen phosphorylase (EC 2.4.1.1) - yeast (Saccharomyces cerevisiae)
N:Alternate names: GPH1 protein; protein P9584.1; protein YPR160w
C:Species: Saccharomyces cerevisiae
C:Date: 23-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 23-Mar-2001
C:Accession: S61144; S29860; A24900; S24010; S32596; S32597; S50939; S40427
R:Stellies, L.
submitted to the EMBL Data Library, June 1995
A:Description: The sequence of S. cerevisiae cosmid 9584.
A:Reference number: S61139
A:Accession: S61144
A:Molecule type: DNA
A:Residues: 1-902 <STE>
A:Cross-references: EMBL:U28371; NID:g849161; PIDN:AAB68057.1; PID:g849168; MIPS:YPR1
R:Fletcherick, R.J.
A:Reference number: S29860
A:Accession: S29860
A:Molecule type: DNA
A:Residues: 1-166, EG', 169-325, 'P', 327-439, 'RR', 442-507, 'V', 509-523, 'I', 525-577, 'E', 5
A:Cross-references: EMBL:X04604; NID:g4172; PIDN:CAA28273.1; PID:g4173
R:HWang, P.K.; Fletcherick, R.J.
Nature 324, 80-84, 1986
A:Title: Convergent and divergent evolution of regulatory sites in eukaryotic phospho
A:Reference number: A24900; MUID:87065036; PMID:3537803
A:Accession: A24900
A:Molecule type: DNA
A:Residues: 12-166, EG', 169-325, 'P', 327-439, 'RR', 442-507, 'V', 509-523, 'I', 525-577, 'E',
A:Cross-references: EMBL:X04604; NID:g4172; PIDN:CAA28273.1; PID:g4173
A>Note: the sequence from Fig. 2 is inconsistent with that from Fig. 1 in having 90-V
J. Mol. Biol. 225, 1027-1034, 1992
A:Title: Purification and crystallization of glycogen phosphorylase from Saccharomyce
A:Reference number: S24010; MUID:92309401; PMID:1613787
A:Accession: S24010
A:Molecule type: protein
A:Residues: 2-8, 'X', 10 <RAT>
R:Lerch, K.; Fischer, E.H.
Biochemistry 14, 2009-2014, 1975
```

A:Title: Amino acid sequence of two functional sites in yeast glycogen phosphorylase.
A:Reference number: S32596; MUID:75146541; PMID:1092346
A:Accession: S32596
A:Molecule type: protein
A:Residues: 13-18,'R',20-24;30-37;737-754 <LER>
R:Becker, J.U.; Wingender-Drissen, R.; Schiltz, E.
Arch. Biochem. Biophys. 225, 667-678, 1983
A:Title: Purification and properties of phosphorylase from baker's yeast.
A:Reference number: S32597; MUID:84022569; PMID:6354094
A:Accession: S32597
A:Molecule type: protein
A:Residues: 'X',38-39,'X',41-48 <BEC>
R:Roemer, T.; Fortin, N.; Bussey, H.
Yeast 10, 1527-1530, 1994
A:Title: DNA sequence analysis of a 10.4 kbp region on the right arm of yeast chromosome
A:Reference number: S50938; MUID:95176711; PMID:7871892
A:Accession: S50938
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-166,'EG',169-325,'P',327-439,'RR',442-507,'V',509-523,'I',525-577,'E',579-
A:Cross-references: GB:I33835; MID:g499698; PIDN:AAB59313.1; PID:g499700
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1994
R:Hudson, J.W.; Golding, G.B.; Crexar, M.M.
J. Mol. Biol. 234, 700-721, 1993
A:Title: Evolution of allosteric control in glycogen phosphorylase.
A:Reference number: S40427; MUID:94076345; PMID:8254668
A:Contents: annotation
C:Genetics:
A:Gene: SGD:GPH1
A:Cross-references: SGD:S0006364; MIPS:YPR160w
A:Map position: 16R
C:Superfamily: phosphorylase
C:Keywords: glycosyltransferase; hexosyltransferase; phosphoprotein; pyridoxal phosphate
F:2-902/Product: glycogen phosphorylase #status experimental <MAT>
F:31/Binding site: phosphate (Thr) (covalent) #status experimental
F:751/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental

Query Match 7.0%; Score 93; DB 2; Length 902;
Best Local Similarity 24.6%; Pred. No. 8.8;
Matches 49; Conservative 29; Mismatches 75; Indels 46; Gaps 10;

QY 45 VPSNVPLK-EVLMKKQKDKVAELENSEFRASFSEKRVYLDTVSGSL--TIYNLTSSDED 101
Db 41 IDTMPLKSRALWNKHQVK-----KENKAEDFQDR-FIDHVVETTLARSLYN--CDDMA 90
QY 102 EYEMSPNITDTM-----KFFLYVLESPLSPTLTCAITNGSIEVQCM 144
Db 91 AYEASMSIRDLNLDVWNKTQQKFTTRDPKRVYVLSLEFLMGRLADNALINNKIE---D 146
QY 145 PEHYNHRLIMYSWDCPMEQCKRNSTSIYFKMENDLPQKIQCTLSNPLNFTSSIIIL-- 202
Db 147 PEDPAASKGK-----PREMIKGLDLDLGFKLEDVLDQEPDAGLNGGLGLRLAACFVDS 199
QY 203 --TTCIPSSGHS-RHYAL 218
Db 200 MATEGIPAWGIGLRYEIGI 218

RESULT 15
I46690
CD80 precursor - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 29-Sep-1999
C:Accession: I46690
R:Isono, T.; Seto, A.
Immunogenetics 42, 217-220, 1995
A:Title: Cloning and sequencing of the rabbit gene encoding T-cell costimulatory molecule
A:Reference number: I46689; MUID:95369849; PMID:7642234
A:Accession: I46690
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-299 <ISO>
A:Cross-references: GB:D49843; MID:g755096; PIDN:BAA08643.1; PID:g755097

C:Superfamily: B-lymphocyte restricted antigen B7

Query Match 7.0%; Score 92.5; DB 2; Length 299;
Best Local Similarity 19.4%; Pred. No. 2.4;
Matches 47; Conservative 38; Mismatches 70; Indels 87; Gaps 12;

QY 48 NVPLKE-----VLWKKQKDKVAELENSEFRASFSEKRVYLDTVSG-SLTIYNLTSSDED 101
Db 52 NISIDELARMRIYQKQDQWVLSISQVEVWPEYKNTFFDIIINLSLMLALRLSDKG 111
QY 102 EY-----EMESPNTIDTWKFFLYVLESPLSPT---LTCA 132
Db 112 TTYCVVQKNENGSPRRHLTSVLSIRADFPVPVSITDI-----GHDPDNNVKRIGCS 162
QY 133 LTNGSIEVQCMIPHYNSHRGLIMYSWDCPMEQCKRNSTSIYFKMENDLPQKIQCTLSNP 192
Db 163 ASGG-----FPEP-----RLAWMEDGEELNAVNTTV---DQDLDTELYSSSEL 203
QY 193 LFNWTS--SIILTTCIPSSGH-----SRHR-----YALIPPLAVITTCIV 231
Db 204 DFNVTNNHSIV---CLIKYGELSVSQIFPWSKPKQEPFIDQLPFWVIIPVSGALVLTAVY 260
QY 232 LY 233
Db 261 LY 262

Search completed: January 28, 2003, 08:40:50
Job time : 16.9816 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 08:38:45 ; Search time 16.3317 Seconds
(without alignments)
634.907 Million cell updates/sec

Title: US-09-730-465-2

Sequence: 1 MVAGSDAGRALGVLSVCLL.....VLYMNGILKDRKPDRTNSN 250

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1326	100.0	250	1 LFA3_HUMAN	P19256 homo sapien
2	131.5	9.9	240	1 CD48_MOUSE	P18181 mus musculus
3	126	9.5	347	1 CD2_HORSE	P37998 equus caball
4	123	9.3	243	1 CD48_HUMAN	P03326 homo sapien
5	115.5	8.7	351	1 CD2_HUMAN	P06729 homo sapien
6	113	8.5	655	1 LY9_HUMAN	Q9hb37 homo sapien
7	106	8.0	490	1 YAM7_SCHPO	Q10062 schizosacch
8	105.5	8.0	773	1 PIGR_RABIT	P01832 oryctolagus
9	103	7.8	240	1 CD48_RAT	P10252 rattus norv
10	97	7.3	344	1 CD2_RAT	P08921 rattus norv
11	92.5	7.0	299	1 CD8Q_RABIT	P42070 oryctolagus
12	91	6.9	654	1 LY9_MOUSE	Q01965 mus musculus
13	90.5	6.8	1113	1 PR3_MOUSE	O70361 mus musculus
14	89	6.7	873	1 LDVR_MOUSE	P98136 mus musculus
15	89	6.7	873	1 LDVR_RAT	P98166 rattus norv
16	88.5	6.7	306	1 CD80_MOUSE	Q00609 mus musculus
17	88	6.6	690	1 AFT1_YEAST	P22149 saccharomyc
18	87	6.6	288	1 CD80_HUMAN	P33681 homo sapien
19	87	6.6	901	1 PHSN_YEAST	P06738 saccharomyc
20	86	6.5	1403	1 YGN1_YEAST	P53127 saccharomyc
21	85.5	6.4	526	1 YHV1_YEAST	P38849 saccharomyc
22	85.5	6.4	767	1 R1RL_HSVSA	Q10137 herpesvirus
23	85.5	6.4	793	1 KPP3_MOUSE	P70188 mus musculus
24	85	6.4	785	1 VD05_VACCC	P21010 vaccinia vi
25	85	6.4	785	1 VD05_VACCV	P04305 vaccinia vi
26	85	6.4	785	1 VD05_VARV	P33069 variola vir
27	85	6.4	1021	1 CONT_RAT	Q63198 rattus norv
28	84.5	6.4	1039	1 ITA4_MOUSE	Q00651 mus musculus
29	84	6.3	310	1 HEMA_RACVI	Q00716 racoon pox
30	84	6.3	583	1 N22M_NEUCR	Q35140 neurospora
31	84	6.3	873	1 LDVR_HUMAN	P98155 homo sapien
32	83.5	6.3	248	1 MYP0_MOUSE	P27573 mus musculus
33	83	6.3	1020	1 CONT_MOUSE	P12960 mus musculus

34	82.5	6.2	248	1 MYP0_RAT	P06907 rattus norv
35	82	6.2	1018	1 CONT_HUMAN	Q12860 homo sapien
36	81.5	6.1	860	1 UL52_HSV6U	P52467 human herpe
37	81.5	6.1	1076	1 HSER_CAVPO	P70106 cavia porce
38	81.5	6.1	1484	1 NME2_HUMAN	Q13224 homo sapien
39	81	6.1	873	1 LDVR_RABIT	P35953 oryctolagus
40	81	6.1	2215	1 SORL_MOUSE	O88307 m sortilin-
41	80	6.0	512	1 V233_FOWPV	Q9J502 fowlpox vir
42	80	6.0	573	1 DPOL_MOUSE	Q9dxe2 mus musculu
43	80	6.0	651	1 VU47_HSV6G	P30005 human herpe
44	79.5	6.0	329	1 CD86_HUMAN	P42081 homo sapien
45	79.5	6.0	353	1 KERA_CHICK	O42235 gallus gall

ALIGNMENTS

RESULT 1
LFA3_HUMAN
ID LFA3_HUMAN STANDARD; PRT; 250 AA.
AC P19256; O96KI9;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lymphocyte function-associated antigen 3 precursor (Ag3) (Antigen CD58) (Surface glycoprotein LFA-3).
GN CD58 OR LFA3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.
RC TISSUE=Erythrocyte;
RX MEDLINE=88009714; PubMed=3309127;
RA Wallner B.P., Frey A.Z., Tizard R., Mattaliano R.J., Hession C., Sanders M.E., Dustin M.L., Springer T.A.;
RT "Primary structure of Lymphocyte function-associated antigen 3 (LFA-3). The ligand of the T lymphocyte CD2 glycoprotein.";
RT J. Exp. Med. 166:923-932(1987).
RL [2]
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=88039074; PubMed=3313052;
RA Seed B.;
RT "An LFA-3 cDNA encodes a phospholipid-linked membrane protein homologous to its receptor CD2.";
RL Nature 329:840-842(1987).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX MEDLINE=98169008; PubMed=9510189;
RA Wallich R., Brenner C., Brand Y., Roux M., Reister M., Meuer S.;
RT "Gene structure, promoter characterization, and basis for alternative mRNA splicing of the human CD58 gene.";
RL J. Immunol. 160:2862-2871(1998).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 29-122.
RX MEDLINE=99218273; PubMed=10200255;
RA Ikenizu S., Sparks L.M., van der Merwe P.A., Harlos K., Stuart D.I., Jones E.V., Davis S.J.;
RT "Crystal structure of the CD2-binding domain of CD58 (lymphocyte function-associated antigen 3) at 1.8-A resolution.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:4289-4294(1999).
CC -!- FUNCTION: LIGAND OF THE T LYMPHOCYTE CD2 GLYCOPROTEIN. THIS INTERACTION IS IMPORTANT IN MEDIATING THYMOCYTE INTERACTIONS WITH THYMIC EPITHELIAL CELLS, ANTIGEN-INDEPENDENT AND -DEPENDENT INTERACTIONS OF T LYMPHOCYTES WITH TARGET CELLS AND ANTIGEN-PRESENTING CELLS AND THE T LYMPHOCYTE ROSETTING WITH ERYTHROCYTES. IN ADDITION, THE LFA-3/CD2 INTERACTION MAY PRIME RESPONSE BY BOTH THE CD2+ AND LFA-3+ CELLS.
CC -!- SUBCELLULAR LOCATION: type I membrane protein (isoform 1) and attached to the membrane by a GPI-anchor (isoform 2).
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.


```
FT CARBOHYD 203 203 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 240 AA; 27383 MW; F3BF6987A9E9C71E CRC64;

Query Match
Best Local Similarity 9.9%; Score 131.5; DB 1; Length 240;
Matches 57; Conservative 36; Mismatches 90; Indels 77; Gaps 11;

QY 13 VLSVCLHLCGFTSCSQOIIYGVYGVNTHVFSNV--PLKEVLWKOKD-KVAELE-N 68
DB 11 VLELLPLGFGGHSIPDINATNGSNVTLKIHKDPGPKYKRTLWLTNOKILEYNN 70

QY 69 SFRAF-SFKNRYVLTGSLTIYNTLTSDEDEYEMESPNTIDT-MKFFLYVLESPLS 126
DB 71 SKTTFSEFGRVYLEENGALHISNRKEDKGYTMYRVLRRTENELKITLEVEDPVK 130

QY 127 PLTLCALNGS-----TEVOCMIPEHNSHRLIMYSDCQMECKRSTSYFKWENDL 181
DB 131 PSTEINKTEASTSCHLRSCVKKOQ-----VDYTW-----YESSGPF 169

QY 182 POK-----IQCTLSNPLNTTSSILATTCIPSSGSHSHRYALIPPL 223
DB 170 PKKPGYVLDLIVTPQNKSTFTYTCQVSNVSSKNDTV-----YFTLPQDL 214

QY 224 A-----VITTCIV 231
DB 215 ARSSGVCTATWLVTTLII 234

RESULT 3
ID CD2_HORSE STANDARD; PRT; 347 AA.
AC P3798;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DE 13-JUN-2002 (Rel. 41, Last annotation update)
DE T-cell surface antigen CD2 precursor.
GN GN
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A., AND 3D-STRUCTURE MODELING.
RC TISSUE=T-cell;
RX MEDLINE=94155904; PubMed=7906650;
RA Taveron A.S., Kydd J.H., Bodian D.L., Jones E.Y., Stuart D.I.,
RA Davis S.J., Butcher G.W.;
RT "Expression cloning of an equine T-lymphocyte glycoprotein CD2 cDNA.
RL Eur. J. Biochem. 219:969-976(1994).
CC -1- FUNCTION: CD2 INTERACTS WITH LYMPOCYTE FUNCTION-ASSOCIATED
CC ANTIGEN (LFA-3) TO MEDIATE ADHESION BETWEEN T-CELLS AND OTHER
CC CELL TYPES. CD2 IS IMPLICATED IN THE TRIGGERING OF T-CELLS, THE
CC CYTOPLASMIC DOMAIN IS IMPLICATED IN THE SIGNALING FUNCTION.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN SPLEEN AND THYMUS.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X69884; CAA49511.1; .
DR PIR: S31578; S31578.
DR PIR: S41638; S41638.
DR HSP: P06729; IGVA.
KW Immunoglobulin domain; T-cell; Glycoprotein; Antigen; Transmembrane;
```

```
KW Cell adhesion; Repeat; Signal.
FT SIGNAL 1 24 BY SIMILARITY.
FT CHAIN 25 347 T-CELL SURFACE ANTIGEN CD2.
FT DOMAIN 25 205 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 206 230 POTENTIAL.
FT DOMAIN 231 347 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 123 IG-LIKE V-TYPE DOMAIN.
FT DOMAIN 124 199 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 283 303 3 X 7 AA TANDEM REPEATS OF H-R-P-[QL]-[AVP]-P-G.
FT REPEAT 283 289 1.
FT REPEAT 290 296 2.
FT REPEAT 297 303 3.
FT DOMAIN 277 340 PRO-RICH.
FT DISULFID 134 199 BY SIMILARITY.
FT DISULFID 141 182 BY SIMILARITY.
FT CARBOHYD 27 27 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 68 68 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 185 185 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 347 AA; 38864 MW; 0529606A2783D5E7 CRC64;

Query Match
Best Local Similarity 9.5%; Score 126; DB 1; Length 347;
Matches 51; Conservative 44; Mismatches 72; Indels 52; Gaps 11;

QY 14 LSVVC-LLHCFGFISCFSQ-----IYGVYGVNTHVFS---NVLPEVLWKQK 60
DB 1 MNLACKLLASFLLIFFSKGAVSKNKTILGALERDINLDIPAFQMSERHVEDIQWSKG 60

QY 61 DKVAELENSEFRASFKNRYVLTGVSGLTIYNTLSSDE-----DEVEMESPNTIDTMK 115
DB 61 TKIAKFKNG---SWTFQDKYIEVVKLTKIKHLRIHEGYKYVDADSDGKNVLEE-T 116

QY 116 FFYLVLESPLSP-----TLTCAITNGS-TEVOCMIPEH--YNSHRGLIMYSWDCPM 163
DB 117 FHLSLLEWMSKPNISWCSCTNTLTCEVTGKTDFELKLYLNGRMIOKSPRKVIVY----- 170

QY 164 EQCKRNSTSYFKMENDLPQKIQCTLSNPLNTTSSIL 202
DB 171 --KRASNQI-----ASFKCTANTVTSESSSVWI 197

RESULT 4
ID CD48_HUMAN STANDARD; PRT; 243 AA.
AC P09326;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE B-lymphocyte activation marker BLAST-1 precursor (BCM1 surface
DE antigen) (Leucocyte antigen MEM-102) (TCT.1) (Antigen CD48).
GN CD48 OR BCM1 OR BLAST1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91153858; PubMed=1999351;
RA Vaughan H.A., Henning M.M., Purcell D.F.J., McKenzie I.F.C.,
RA Sandrin M.S.;
RT "The isolation of cDNA clones for CD48."
RL Immunogenetics 33:113-117(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91153857; PubMed=1999350;
RA Korinek V., Stefanova I., Angelisova P., Hilbert I., Horejsi V.;
RT "The human leucocyte antigen CD48 (MEM-102) is closely related to the
RT activation marker Blast-1."
RL Immunogenetics 33:108-112(1991).
```

RN [3] SEQUENCE FROM N.A.
 RX MEDLINE=88111558; PubMed=2828034;
 RA Staunton D.E., Thorley-Lawson D.A.;
 RT "Molecular cloning of the lymphocyte activation marker Blast-1.";
 RL EMBO J. 6:3695-3701(1987).
 RN [4]
 RP SEQUENCE OF 1-27 FROM N.A.
 RX MEDLINE=91141511; PubMed=1847502;
 RA Fisher R.C., Thorley-Lawson D.A.;
 RT "Characterization of the Epstein-Barr virus-inducible gene encoding
 the human leukocyte adhesion and activation antigen BLAST-1 (CD48).";
 RL Mol. Cell. Biol. 11:1614-1623(1991).
 RN [5]
 RP PARTIAL SEQUENCE.
 RX MEDLINE=91237281; PubMed=1827826;
 RA del Porto P., Mami-Chouaib F., Bruneau J.-M., Jitsukawa S., Dumas J.,
 RA Harnois M., Hercend T.;
 RT "TCT-1, a target molecule for gamma/delta T cells, is encoded by an
 immunoglobulin superfamily gene (blast-1) located in the CD1 region
 of human chromosome 1.";
 RL J. Exp. Med. 173:1339-1344(1991).
 CC -|- FUNCTION: LIGAND FOR CD2. MIGHT FACILITATE INTERACTION BETWEEN
 CC ACTIVATED LYMPHOCYTES. PROBABLY INVOLVED IN REGULATING T-CELL
 CC ACTIVATION.
 CC -|- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC -|- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC -|- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
 CC -|- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
 CC -|- DATABASE: NAME=PROW; NOTE=CD guide CD48 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd48.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M59904; AAA62834.1; -;
 DR EMBL; X06341; CAA29647.1; -;
 DR EMBL; M37766; AAA36211.1; -;
 DR EMBL; M63911; AAA35602.1; -;
 DR PIR; A27236; A27236.
 DR PIR; A53244; A53244.
 DR Genew; HGNC:1683; CD48.
 DR MIM; 109530; -;
 DR InterPro; IPR003598; Ig.
 DR InterPro; IPR003006; Ig_MHC.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00409; IG; 1.
 KW Immunoglobulin domain; B-cell; Glycoprotein; GPI-anchor; T-cell;
 KW Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 220
 FT PROPEP 221 243
 FT LIPTD 220 220
 FT DISULFID 154 196
 FT CARBOHYD 40 40
 FT CARBOHYD 44 44
 FT CARBOHYD 104 104
 FT CARBOHYD 162 162
 FT CARBOHYD 189 189
 FT CONFLICT 2 2
 FT CONFLICT 66 66
 FT CONFLICT I -> N (IN REF. 3).
 SQ SEQUENCE 243 AA; 27683 MW; 9E46D76FC36A512C CRC64;
 Query Watch 9.3%; Score 123; DB 1; Length 243;
 Best Local Similarity 27.0%; Pred. No. 0.0017;
 Matches 50; Conservative 32; Mismatches 79; Indels 24; Gaps 9;
 QY 36 VYVG-VNTHVPSNVP--LKEVLWKQKD-KVAELENSEPRF--SSFKNRVYLDTVSGSL 90

Db 35 VVSGNVTNLINISESLPENYKQLTWFTYTDQKIVEDWSRKSIFYESKFKGRVRLDQSGAL 94
 QY 91 TIYNLTSSDEYEMESPNT---DTMKFFLYVLESPLTLCAL-----TNGSIEVQC 142
 Db 95 YISKVQKEDNSTYIMRVLKKTGNEQEWKIKQLQVDPKPKVRIKTEKDNDONCVLKLSC 154
 QY 143 MIP-EHYNSHRGLIMYSW--DCPMEOCKRNSTSIYFKMENDLPQKIOCTLSNPLFNNTS 198
 Db 155 VIPGESVN-----YIYGDKRPFKELQNSVLETTLMPHNSRCYTCOVNSVSSKNG 207
 QY 199 SIILT 203
 Db 208 TVCLS 212
 RESULT 5
 CD2_HUMAN
 ID CD2_HUMAN STANDARD; PRT; 351 AA.
 AC P06729;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE T-cell surface antigen CD2 precursor (T-cell surface antigen
 DE TIL/Leu-5) (LFA-2) (LFA-3 receptor) (Erythrocyte receptor) (Rosette
 DE receptor).
 GN CD2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8814486; PubMed=2894031;
 RA Diamond D.J., Clayton L.K., Sayre P.H., Reinherz E.L.;
 RT "Exon-intron organization and sequence comparison of human and murine
 RT T11 (CD2) genes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:1615-1619(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87204137; PubMed=2437578;
 RA Seed B., Aruffo A.;
 RT "Molecular cloning of the CD2 antigen, the T-cell erythrocyte
 RT receptor, by a rapid immunoselection procedure.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:3365-3369(1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87041523; PubMed=3490670;
 RA Sewell W.A., Brown M.H., Dunne J., Owen M.J., Crumpton M.J.;
 RT "Molecular cloning of the human T-lymphocyte surface CD2 (T11
 RT antigen).";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:8718-8722(1986).
 RN [4]
 RP REVISIONS.
 RA Sewell W.A., Brown M.H., Dunne J., Owen M.J., Crumpton M.J.;
 RL Proc. Natl. Acad. Sci. U.S.A. 84:7256-7256(1987).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87204243; PubMed=2883656;
 RA Sayre P.H., Chang H.-C., Hussey R.E., Brown N.R., Richardson N.E.,
 RA Spagnoli G., Clayton L.K., Reinherz E.L.;
 RT "Molecular cloning and expression of T11 cDNAs reveal a receptor-like
 RT structure on human T lymphocytes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:2941-2945(1987).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89005055; PubMed=2901953;
 RA Lang G., Motton D., Owen M.J., Sewell W.A., Brown M.H., Mason D.Y.,
 RA Crumpton M.J., Kioussis D.;
 RT "The structure of the human CD2 gene and its expression in transgenic
 RT mice.";
 RL EMBO J. 7:1675-1682(1988).
 RN [7]

CC -!- SIMILARITY: TO PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) AND TO
 CC YEAST YER014W.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Z68136; CAA92235.1; -
 CC InterPro: IPR002937; Amino_oxidase.
 CC InterPro: IPR004572; Proto_IX_ox.
 CC Pfam: PF01593; Amino_oxidase; 1.
 CC TIGRFAMs: TIGR00562; proto_IX_ox; 1.
 CC KW Hypothetical protein.
 CC SEQUENCE 490 AA; 53461 MW; CB679F6A437547AA CRC64;

Query Match 8.0%; Score 106; DB 1; Length 490;
 Best Local Similarity 23.9%; Pred. No. 0.12; Mismatches 58; Gaps 14;
 Matches 62; Conservative 43;
 QY 23 FGFISCFSGQIQGVGVNVTVPVSNVPLKEVLKKQK-----DKVAELENSEFRA 73
 Db 191 FGFLAKTERK-----YGNITGLIRALLAREILSPAELKAAALAEKPAELSNLS-MKS 244
 QY 74 FSPFKRVYLDVSGSLIYNTSDDEYEMESPNTITDMKFFLYLVLESPLTLCAL 133
 Db 245 TSMFAKKEGIEI-----TLSIADELK-RMPNV-----NKPAKTLVPHK 286
 QY 134 TNGSIEVQCMPEHY---NSHRLIMYSWDCPMEQCKRNS---TSIYFKMENDLPQK--- 184
 Db 287 TQSLVDVNGQAVEYVVFANSRNLENLI-SCPKNMETPTSSVYVWVYKDPNVLPIRGFG 345
 QY 185 ---IQCTLSP-----LF-----NTTSIIITLCIPSSGHSRHRVYALIPPLAVITCI 230
 Db 346 LLIPSCIPNPNHVLGIVDFEQNPENGSKVIVMVGSAITKNT-SLIPTNPEAVNNA 404
 QY 231 VLYNGLIKCDKRRPQRTNS 249
 Db 405 LKALQHTLTKISSKPTLTNA 423

RESULT 8

ID PIGR_RABIT STANDARD; PRT; 773 AA.
 AC P01832;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Polymetric-immunoglobulin receptor precursor (Poly-IG receptor) (PIGR)
 DE [Contains: Secretory component].
 GN PIGR.
 OS Oryctolagus cuniculus (Rabbit).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 CC NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84142246; PubMed=6322002;
 RA Mostov K.E., Friedlander M., Blobel G.;
 RT "The receptor for transepithelial transport of IgA and IgM contains
 RT multiple immunoglobulin-like domains.";
 RL Nature 308:37-43(1984).
 RN [2]
 RP SEQUENCE OF 87-114 AND 410-428.
 RX MEDLINE=88228032; PubMed=3131339;
 RA Frutiger S., Hughes G.J., Hanly W.C., Jaton J.-C.;
 RT "Rabbit secretory components of different allotypes vary in their
 RT carbohydrate content and their sites of N-linked glycosylation.";
 RL J. Biol. Chem. 263:8120-8125(1988).
 CC -!- FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE

CC BASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN
 CC TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE.
 CC DURING THIS PROCESS A CLEAVAGE OCCURS THAT SEPARATE THE
 CC EXTRACELLULAR (KNOWN AS THE SECRETORY COMPONENT) FROM THE
 CC TRANSMEMBRANE SEGMENT.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ALSO SECRETED.
 CC -!- POLYMORPHISM: THE SEQUENCE SHOWN IS THAT OF ALLOTYPES T61.
 CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE V-TYPE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X00412; CAA25118.1; -
 CC PIR: A02111; ORRG.
 CC PIR: A28077; A28077.
 CC InterPro: IPR003599; Ig.
 CC InterPro: IPR003006; Ig_MHC.
 CC InterPro: IPR003600; Ig_Like.
 CC Pfam: PF00047; Ig; 5.
 CC SMART: SM00409; Ig; 5.
 CC SMART: SM00410; Ig_Like; 1.
 CC Immunoglobulin domain; Repeat; Transmembrane; Glycoprotein; Signal;
 CC Polymorphism.
 CC SIGNAL 1 18
 FT CHAIN 19 773 POLYMERIC-IMMUNOGLOBULIN RECEPTOR.
 FT CHAIN 19 615 SECRETORY COMPONENT.
 FT DOMAIN 19 647 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 648 670 POTENTIAL.
 FT DOMAIN 671 773 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 39 122 IG-LIKE V-TYPE DOMAIN 1.
 FT DOMAIN 148 232 IG-LIKE V-TYPE DOMAIN 2.
 FT DOMAIN 253 331 IG-LIKE V-TYPE DOMAIN 3.
 FT DOMAIN 362 445 IG-LIKE V-TYPE DOMAIN 4.
 FT DOMAIN 471 545 IG-LIKE V-TYPE DOMAIN 5.
 FT DISULFID 46 115 POTENTIAL.
 FT DISULFID 155 225 POTENTIAL.
 FT DISULFID 260 324 POTENTIAL.
 FT DISULFID 369 438 POTENTIAL.
 FT DISULFID 478 538 POTENTIAL.
 FT CARBOHYD 88 108 N-LINKED (GLCNAC. .); IN ALLOTYPES T61.
 FT CARBOHYD 108 108 N-LINKED (GLCNAC. .); IN ALLOTYPES T62
 FT CARBOHYD 108 108 (PARTIAL) AND T63.
 FT CARBOHYD 418 418 N-LINKED (GLCNAC. .).
 FT VARIANT 88 88 K -> N (IN ALLOTYPES T61).
 FT VARIANT 94 94 D -> E (IN ALLOTYPES T61).
 FT VARIANT 101 108 TVDOLTON -> YLNRLSOS (IN ALLOTYPES T61).
 FT VARIANT 110 110 S -> T (IN ALLOTYPES T63).
 SQ SEQUENCE 773 AA; 83886 MW; DF2C44D2F1193C65 CRC64;
 Query Match 8.0%; Score 105.5; DB 1; Length 773;
 Best Local Similarity 22.0%; Pred. No. 0.23;
 Matches 47; Conservative 39; Mismatches 79; Indels 49; Gaps 11;
 QY 39 GNVTHVP-----SNVPLKEVLKKQKDKVAELENSEFRAESSFKRVYLDVSG-----S 89
 Db 363 GSVTHRCYPNKRSDSHLQYLWEGSGRHLVDSGGLVQKQYTGRLALFEPPGNGTFS 422
 QY 90 LTIYNIUTSSDEDEY---EMESPNTITDMKFFLYLVLESPLTLP---CALTFNGSIEVQCM 144
 Db 423 VVLNQLTADEGCFYWCVSDDDESILTSVK--LQIVDGEPSPTIDKFTAVQGEPEITCHF 480
 QY 145 P-EHYNHSHRGLIMYSWDCPMEQCKRNSIYFKMENDLPQKIQCTLSNPLNTSSIIIT 203
 Db 481 PKYFSSEK-----YW-----CKWNHGCC-----EDLPTKL---SSSGDLVKNNNLVIT 522
 QY 204 T-----CIPSSGHSRHRVYALIPPL 223

Db 523 LTLDVSDEGWYCGAKGDBEVEAAVRVEL 556

RESULT 9

CD48_RAT

AC CD48_RAT STANDARD; PRT; 240 AA.

AC P10252;

DT 01-MAR-1989 (Rel. 10, Created)

DT 01-MAR-1989 (Rel. 10, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE MRC OX-45 surface antigen precursor (BCM1 surface antigen) (BLAST-1)

DE (CD48).

DE CD48 OR BCM-1.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP MEDLINE=89030603; PubMed=3181129;

RA Killeen N., Moessner R., Arvieux J., Willis A., Williams A.F.;

RX "The MRC OX-45 antigen of rat leukocytes and endothelium is in a subset of the immunoglobulin superfamily with CD2, LFA-3 and carcinoembryonic antigens.";

RT EMBO J. 7:3087-3091(1988).

RL

CC -1- FUNCTION: LIGAND FOR CD2. MIGHT FACILITATE INTERACTION BETWEEN ACTIVATED LYMPHOCYTES. PROBABLY INVOLVED IN REGULATING T-CELL ACTIVATION.

CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.

CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.

CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

DR EMBL; X13016; CAA31438.1; -

DR PIR; S01299; S01299.

DR GlycoSuiteDB; P10252; -

DR InterPro; IPR003599; Ig.

DR Pfam; PF00047; Ig; 1.

DR SMART; SM00409; Ig; 1.

KW Immunoglobulin domain; T-cell; Antigen; Glycoprotein; GPI-anchor; Signal.

FT SIGNAL. 1 22

FT CHAIN 23 217 MRC OX-45 SURFACE ANTIGEN.

FT PROPEP 218 240 REMOVED IN MATURE FORM.

FT LIPID 217 217 GPI-ANCHOR.

FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 203 203 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 240 AA; 27680 MW; 42CBFB9947A0E437 CRC64;

Query Match 7.88; Score 103; DB 1; Length 240;

Best Local Similarity 21.9%; Pred. NO. 0.088;

Matches 55; Conservative 43; Mismatches 99; Indels 54; Gaps 14;

Qy 13 VLSVCLLCRCGFCISCFSGQIYGVVGNVTFHVPNSVPLKEVLKKOK-----DKVA 64

Db 11 ILESLLSLVYGFQDSQVNVNVAITGSNVTLI-----LKHPLASYORLTWLTHTNQKIL 65.

Qy , 65 E-LENSEFRASS-FKNRYVLDYTSGLSTIYNLTSSDEDEYEMESPNT-DTMKFFLYVL 121

Db 66 EYFNGKKTVFESFKDRVLDKTGALRIYVNSKDRGDIYMRMLHETEDQWKITMEVY 125

Qy 122 ESLSPPTL----TCALTING-SIEVOCMIPHEYNHSHRGLIMYSW---DCPMEQ----- 165

Db 126 DLVSKPAIKTEKTKNLTDSCHLRLSCKVED-----QG-VDTWYEDSGFPORNPGYVLE 179

Qy 166 ---CKRNSTSIYFKMENDLPQKIOTLNPLFNFTSSI-ILTTCTIPSSGHSRHYAL-IP 220

Db 180 ITITPHNKSTFY-----TCQVSNPVSSENDTLYPIPPCTLARSSGVHIAAWLV 228

Qy 221 IPLAVITTCIV 231

Db 229 VTLSIIPSILL 239

RESULT 10

CD2_RAT

ID CD2_RAT STANDARD; PRT; 344 AA.

AC P08921;

DT 01-NOV-1988 (Rel. 09, Created)

DT 01-NOV-1988 (Rel. 09, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE T-cell surface antigen CD2 precursor (T-cell surface antigen T11/Leu-5) (LFA-2) (LFA-3 receptor) (OX-34 antigen).

DE CD2.

GN

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE OF 42-344 FROM N.A., AND PARTIAL SEQUENCE.

RC STRAIN=AO;

RX MEDLINE=87113973; PubMed=3102667;

RA Williams A.F., Barclay A.N., Clark S.J., Paterson D.J., Willis A.C.;

RT "Similarities in sequences and cellular expression between rat CD2 and CD4 antigens.";

RL J. Exp. Med. 165:368-380(1987).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=AO;

RA Barclay A.N., Williams A.F.;

RL Submitted (MAY-1987) to the EMBL/GenBank/DBJ databases.

RN [3]

RP IMPORTANCE OF C-TERMINAL IN SIGNALING.

RX MEDLINE=88327862; PubMed=2901293;

RA He Q., Beyers A.D., Barclay A.N., Williams A.F.;

RT "A role in transmembrane signaling for the cytoplasmic domain of the CD2 T lymphocyte surface antigen.";

RL Cell 54:979-984(1988).

RN [4]

RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 23-198.

RX MEDLINE=93063364; PubMed=1279440;

RA Jones E.Y., Davis S.J., Williams A.F., Harlos K., Stuart D.I.;

RT "Crystal structure at 2.8-A resolution of a soluble form of the cell adhesion molecule CD2.";

RL Nature 360:232-239(1992).

RN [5]

RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 23-121.

RX MEDLINE=95365363; PubMed=7638192;

RA Murray A.J., Lewis S.J., Barclay A.N., Brady R.L.;

RT "One sequence, two folds: a metastable structure of CD2.";

RL Proc. Natl. Acad. Sci. U.S.A. 92:7337-7341(1995).

RN [6]

RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 23-121.

RX MEDLINE=98400497; PubMed=9731771;

RA Murray A.J., Head J.G., Barker J.J., Brady R.L.;

RT "Engineering an intertwined form of CD2 for stability and assembly.";

RL Nat. Struct. Biol. 5:778-782(1998).

RN [7]

RP STRUCTURE BY NMR OF 23-121.

RX MEDLINE=92049704; PubMed=1682812;

RA Driscoll P.C., Cyster J.G., Campbell I.D., Williams A.F.;

RT "Structure of domain 1 of rat T lymphocyte CD2 antigen.";

RL Nature 353:762-765(1991).

CC -1- FUNCTION: CD2 INTERACTS WITH LYMPHOCYTE FUNCTION-ASSOCIATED

CC ANTIGEN (LFA-3) AND OX-45/BCM-1 TO MEDIATE ADHESION BETWEEN T
 CC CELLS AND OTHER CELL TYPES. CD2 IS IMPLICATED IN THE TRIGGERING
 CC OF T-CELLS. THE CYTOPLASMIC DOMAIN IS IMPLICATED IN THE
 CC SIGNALING FUNCTION.
 CC
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
 CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X05111; CAA28757.1; -
 CC PIR: A33071; PRWYC2.
 CC PIR: A32346; A32346.
 CC PDB: 1CDC; 15-SEP-95.
 CC PDB: 1HNG; 07-FEB-95.
 CC PDB: 1A64; 27-MAY-98.
 CC PDB: 1A6P; 17-JUN-98.
 CC PDB: 1A7B; 17-JUN-98.
 CC InterPro: IPR003600; Ig_Like.
 CC SMART: SM00410; Ig_Like; 1.
 CC Immunoglobulin domain; T-cell; Glycoprotein; Antigen; Transmembrane;
 KW Cell adhesion; Repeat; Signal; 3b-structure.
 FT SIGNAL 1 22
 FT CHAIN 23 344 T-CELL SURFACE ANTIGEN CD2.
 FT DOMAIN 23 202 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 203 228 POTENTIAL.
 FT DOMAIN 229 344 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 23 121 IG-LIKE V-TYPE DOMAIN.
 FT DOMAIN 122 202 IG-LIKE C2-TYPE DOMAIN.
 FT DOMAIN 277 343 PRO-RICH.
 FT DISULFID 132 196
 FT DISULFID 139 179
 FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .).
 SQ SEQUENCE 344 AA; 38414 MW; 41BAED392CE16356 CRC64;
 Query Match 7.3%; Score 97; DB 1; Length 344;
 Best Local Similarity 27.9%; Pred. No. 0.45;
 Matches 50; Conservative 26; Mismatches 69; Indels 34; Gaps 13;
 QY 35 GVYGVNVTFRVPSNVP-----LKEVLMKKQDKVAELENSEFRASFKNRVYLDTVS 87
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 Db 26 GTVWGALGHGILNIPNFQMTDDIDEVWERSGTLVAEFKR-KMKPF--LKSGAFETLAN 82
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 QY 88 GSLLIYNLTSSDEDEYEME--SPNITDTM--KFFLYVLSLSP-----TLTCAIT 134
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 Db 83 GDLKIKNLTRDDSGTNYVTVYSTNGTRILDKALDRLKLEWVSKPMIYWECSNATLTCEVL 142
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 QY 135 NGSEVQCMT---PEHNSHRLGIM-YSWD---CPMEQCKR-NSTSIYFKME-NDLPQK 184
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 Db 143 EGT-DVELKLYQCKEHLRSRQKMTSYQWTLNRAPE-KCKAVNRVYSESEMEVYVNCPEK 199
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 RESULT 11
 CD80_RABIT STANDARD; PRT; 299 AA.
 ID CD80_RABIT
 AC P42070;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE T lymphocyte activation antigen CD80 precursor (Activation B7-1
 DE antigen).
 GN CD80.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-B/J X CHBB;HM;
 RX MEDLINE=95369849; PubMed=7642234;
 RA Isono T., Seto A.;
 RT "Cloning and sequencing of the rabbit gene encoding T-cell
 RT costimulatory molecules";
 RL Immunogenetics 42:217-220(1995).
 CC -!- FUNCTION: INVOLVED IN THE COSTIMULATORY SIGNAL ESSENTIAL FOR T
 CC LYMPHOCYTES ACTIVATION. T CELL PROLIFERATION AND CYTOKINE
 CC PRODUCTION IS INDUCED BY THE BINDING OF CD28 OR CTLA-4 TO THIS
 CC RECEPTOR.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
 CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: D49843; BAA08643.1; -
 CC HSSP: P33681; IDR9.
 CC InterPro: IPR003599; Ig.
 CC InterPro: IPR003006; Ig_MHC.
 CC InterPro: IPR003600; Ig_Like.
 CC Pfam: PF00047; Ig_1.
 CC SMART: SM00409; Ig_1.
 CC SMART: SM00410; Ig_Like; 1.
 KW Immunoglobulin domain; T-cell; Glycoprotein; Signal; Transmembrane;
 FT SIGNAL 1 32 POTENTIAL.
 FT CHAIN 33 299 T LYMPHOCYTE ACTIVATION ANTIGEN CD80.
 FT DOMAIN 33 243 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 244 264 POTENTIAL.
 FT DOMAIN 265 299 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 42 122 IG-LIKE V-TYPE DOMAIN.
 FT DOMAIN 154 222 IG-LIKE C2-TYPE DOMAIN.
 FT DISULFID 49 115 POTENTIAL.
 FT DISULFID 161 215 POTENTIAL.
 FT CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 185 185 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 206 206 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 210 210 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 299 AA; 33513 MW; 674423E5CC91DE0 CRC64;
 Query Match 7.0%; Score 92.5; DB 1; Length 299;
 Best Local Similarity 19.4%; Pred. No. 0.91;
 Matches 47; Conservative 38; Mismatches 70; Indels 87; Gaps 12;
 QY 48 NYPLEK-----VLKKQDKVAELENSEFRASFKNRVYLDTVSG-SLTIYNLTSSDED 101
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 Db 52 NSIDELARMRIYWKQDQWVLSIISQGVVEWPEYKNTFFDIINNLSMLALRLSDKG 111
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 QY 102 EY-----EMESPNTIDTMKFFLYVLSLSPST---LTCA 132
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 Db 112 TVTCVQVKNENGSPFREHLTSVTLSIRADFPVPSITDI-----GHPDPNVKIRGS 162
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 QY 133 LINGSIEVQCMPEHNSHRLGIMYSWDCPEMCKENSTSIYFKMENDLPQKLTCLSNP 192
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 Db 163 ASGG-----FPPE-----FLAWMEDGELNAVNTIV-----DQDQDTLYSVSSEL 203
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 QY 193 LFNWTS--SIILTTCTPSSGH-----SRHR-----YALIPLAIVITTCIV 231
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

DB 204 DFNVTNHSIV---CLIKYGELSVSQIFPWSKPKQEPIDQLPFWIIPVSGALVLTAVV 260
QY 232 LY 233
DB 261 LY 262

RESULT 12
LY9_MOUSE
ID LY9_MOUSE STANDARD; PRT; 654 AA.
AC Q01965; Q9ES35; Q9ES36; Q9ES29;
DT 01-JUN-1994 (Rel. 29, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE T-lymphocyte surface antigen Ly-9 precursor (Lymphocyte antigen 9)
DE (Cell-surface molecule Ly-9).
GN LY9 OR LY-9.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A., AND POLYMORPHISM.
RX STRAIN=129/SV, BALB/c, and C57BL/6; TISSUE=Spleen;
RX MEDLINE=20424510; PubMed=10970093;
RA Tovar V., de la Fuente M.A., Pizcueta P., Bosch J., Engel P.;
RA "Gene structure of the mouse leukocyte cell surface molecule Ly9.";
RL Immunogenetics 51:788-793(2000).
RN [2]
RP SEQUENCE OF 22-654 FROM N.A., AND SEQUENCE OF 48-59.
RX MEDLINE=92373005; PubMed=1506686;
RA Sandrin M.S., Gumley T.P., Henning M.M., Vaughan H.A., Gopez L.J.,
RA Trapani J.A., McKenzie I.F.C.;
RL J. Immunol. 149:1636-1641(1992).
CC "- Isolation and characterization of cDNA clones for mouse Ly-9.";
CC "- FUNCTION: May participate in adhesion reactions between T
CC lymphocytes and accessory cells by homophilic interaction.
CC "- SUBCELLULAR LOCATION: Type I membrane protein.
CC "- TISSUE SPECIFICITY: LYMPHOCYTES.
CC "- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC "- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC "- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE V-TYPE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR EMBL; AF244131; AAG14997.1; -;
DR EMBL; AF244130; AAG14996.1; -;
DR EMBL; AF246701; AAG13268.2; -;
DR EMBL; AF245117; AAG13268.2; JOINED.
DR EMBL; AF245506; AAG13268.2; JOINED.
DR EMBL; AF245118; AAG13268.2; JOINED.
DR EMBL; AF245507; AAG13268.2; JOINED.
DR EMBL; AF245508; AAG13268.2; JOINED.
DR EMBL; AF245509; AAG13268.2; JOINED.
DR EMBL; AF245510; AAG13268.2; JOINED.
DR EMBL; AF246699; AAG13268.2; JOINED.
DR EMBL; AF246700; AAG13268.2; JOINED.
DR EMBL; M84412; ARA39468.1; -;
DR PIR; A46500; A46500.
DR HSP; P08921; 1HNG.
DR MGD; MGI:96885; Ly9.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00409; Ig; 2.
DR SMART; SM00410; Ig_like; 3.

KW Cell adhesion; Antigen; Signal; Transmembrane; Glycoprotein;
KW Immunoglobulin domain; Repeat; Polymorphism.
FT SIGNAL 1 47
FT CHAIN 48 654
FT DOMAIN 48 453
FT TRANSMEM 454 474
FT DOMAIN 475 654
FT DOMAIN 48 158
FT DOMAIN 163 249
FT DOMAIN 250 382
FT DOMAIN 363 453
FT DISULFID 172 242
FT DISULFID 178 222
FT DISULFID 376 445
FT DISULFID 382 426
FT CARBOHYD 68 68
FT CARBOHYD 120 120
FT CARBOHYD 231 231
FT CARBOHYD 284 284
FT CARBOHYD 390 390
FT CARBOHYD 412 412
FT CARBOHYD 423 423
FT CARBOHYD 434 434
FT VARIANT 10 10
FT VARIANT 14 14
FT VARIANT 79 79
FT VARIANT 91 91
FT VARIANT 130 130
FT VARIANT 139 139
FT VARIANT 362 362
FT VARIANT 366 366
FT VARIANT 377 377
FT VARIANT 550 550
FT VARIANT 592 592
FT CONFLICT 283 283
FT CONFLICT 499 499
FT CONFLICT 560 560
FT CONFLICT 647 654
SQ SEQUENCE 654 AA; 73142 MW; 1CBBB9708AE8EE7 CRC64;
Query Match 6.9%; Score 91; DB 1; Length 654;
Best Local Similarity 24.4%; Pred. No. 3.2;
Matches 54; Conservative 29; Mismatches 78; Indels 60; Gaps 13;
QY 33 IYGVVYGNVTF-----HVPNNVPLKVLWKQDKVAELNSENSEFRAFSSFNRV 81
DB 55 ISGMGGSVTFSLNISKDAEIEHIWNCPPKALALFYKKDITLD-----KYNGRL 107
QY 82 YLDTVSGSLTIYMLTSSDEDEY--EMESPNI--TDTMKFLVYVLESL-----PS 126
DB 108 KVSDEGYSLYMSNLTKSDSGSYHAQINQKNVLTNNKEFTLHIEKLQRPQIIVESVTPS 167
QY 127 PTLTCAITNGSIEVQCMIEPHYNSHRLIMYSW---DCPM-----EOCKRNSTSYFKME 178
DB 168 DTDSCTFT-----LICTVKGTKDS-----VQYSWTRDTHLNTYDGSHTLRYVSQSV---CD 215
QY 179 NDLPOKTOCTLSNPLNPTTSS--IILTTCTIPSSGHSRHYA 217
DB 216 PDLP--YTCKAWNPVSQNSQPVRIWQFC---TGASRRKTA 251
RESULT 13
PER3_MOUSE
ID PER3_MOUSE STANDARD; PRT; 1113 AA.
AC O70361;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Period circadian protein 3 (mPER3).
GN PER3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=96318231; PubMed=9655499;
RA Zyika M.J., Shearman L.P., Weaver D.R., Reppert S.M.;
RT "Three period homologs in mammals: differential light responses in the
RT suprachiasmatic circadian clock and oscillatin transcripts outside of
RT brain.";
RL Neuron 20:1103-1110(1998).
CC
CC -|- FUNCTION: CIRCADIAN REGULATOR THAT MAY ACT AS A TRANSCRIPTION
CC FACTOR. BEHAVES AS A NEGATIVE ELEMENT IN CIRCADIAN TRANSCRIPTIONAL
CC LOOP. DOES NOT APPEAR TO BIND DNA, SUGGESTING INDIRECT
CC TRANSCRIPTIONAL INHIBITION (BY SIMILARITY). EXPRESSION OSCILLATES
CC IN THE SUPRACHIASMATIC NUCLEI (SCN) AND EYES. THE EXPRESSION
CC RHYTHMS APPEAR TO ORIGINATE FROM RETINA.
CC -|- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -|- TISSUE SPECIFICITY: WIDELY EXPRESSED. EXPRESSED IN HEART, BRAIN,
CC LUNG, LIVER, SKELETAL MUSCLE, TESTIS, AND AT LOW LEVEL IN THE
CC SPLEEN AND KIDNEY. IN BRAIN, MAINLY FOUND IN THE SCN, HIPPOCAMPUS,
CC PIRIFORM CORTEX, AND CEREBELLUM. LOWER LEVEL OF EXPRESSION IN THE
CC NEOCORTEX. EXPRESSION EXHIBITS SYNCHRONOUS OSCILLATIONS IN LIVER,
CC SKELETAL MUSCLE AND TESTIS.
CC -|- INDUCTION: NOT ACUTELY INDUCED BY LIGHT IN SCN DURING SUBJECTIVE
CC NIGHT.
CC -|- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
CC -|- SIMILARITY: CONTAINS 2 PAS (PER-ARNT-SIM) DIMERIZATION DOMAINS.
CC -|- SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: AF050182; AAC40147.1; -.
CC MGD: MGI:1277134; Per3.
CC InterPro: IPR001610; PAC.
CC InterPro: IPR000014; PAS_domain.
CC Pfam: PF00989; PAS; 1.
CC SMART: SM00086; PAC; 1.
CC SMART: SM00091; PAS; 1.
CC PROSITE: PSS0112; PAS; 1.
CC Transcription regulation; Nuclear protein; Repeat; Biological rhythms.
KW DOMAIN 52 86 HELIX-LOOP-HELIX MOTIF.
FT DOMAIN 121 187 PAS 1.
FT DOMAIN 256 326 PAS 2.
FT DOMAIN 336 376 PAC.
FT DOMAIN 562 565 POLY-SER.
SQ SEQUENCE 1113 AA; 120939 MW; 8121E235D100A627 CRC64;

Query Match 6.8%; Score 90.5; DB 1; Length 1113;
Best Local Similarity 25.0%; Pred. No. 6.8;
Matches 49; Conservative 28; Mismatches 84; Indels 35; Gaps 9;

QY 31 QQIVGVYGVNVTFFHPSNVPLKKEVLKKQDKVAEL-----ENSEFRAFSPFKNRVY 82
Db ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
456 QQVYASV--NKIKNVGQQLYI-ESMARSSVPKVAETCVPQGGQDKQFSSOITLKNKST 512
QY 83 LDTVSGSLTYINTLSDDEYEMESPNTIDMKFFL--VYLESPLSPITLTCALINGSTEV 140
Db ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
513 TDTGSGG----NL-QEQEPSSSYQQQMNCIDSVIRYLSYSLPAKPKKCIKNTSSSSSE 567
QY 141 QCMPEHYNHSHRGLIMYSWDCPMEQC-----KRNSTSYFKMKNDLPKQICTLSNPLFNT 196
Db ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
568 AKPIPEVDSQSR-----DTEQLLDIRKQETTG-----PSDIDGGGAARTLSTAALSV 614
QY 197 TSSILLTCTIPSSGHS 212
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
615 ASGISQCSCSSTSGHA 630

```

```

RESULT 14
LDVR_MOUSE
ID LDVR_MOUSE STANDARD; PRT; 873 AA.
AC P98156; Q64022;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Very low-density lipoprotein receptor precursor (VLDL receptor).
GN VLDLR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Heart;
RX MEDLINE=95010090; PubMed=7925422;
RA Oka K., Ishimura-Oka K., Chu M.J., Sullivan M., Krushkal J.,
RA Li W.H., Chan L.;
RT "Mouse very-low-density-lipoprotein receptor (VLDLR) cDNA cloning,
RT tissue-specific expression and evolutionary relationship with the
RT low-density-lipoprotein receptor.";
RL Eur. J. Biochem. 224:975-982(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=94283285; PubMed=8013374;
RA Gafvels M.E., Paavola L.G., Boyd C.O., Nolan P.M., Wittmaack F.,
RA Chawla A., Lazar M.A., Bucan M., Angelin B.O., Strauss J.F.;
RT "Cloning of a complementary deoxyribonucleic acid encoding the murine
RT homolog of the very low density lipoprotein/apolipoprotein-E
RT receptor: expression pattern and assignment of the gene to mouse
RT chromosome 19.";
RL Endocrinology 135:387-394(1994).
RN [3]
RP SEQUENCE OF 204-262 FROM N.A.
RX MEDLINE=95003355; PubMed=7919660;
RA Nagdett J.K., Mu J.L.;
RT "The mouse very low density lipoprotein receptor (Vldlr) gene maps to
RT chromosome 19.";
RL Mamm. Genome 5:453-455(1994).
RN [4]
RP BINDING TO REELIN.
RX MEDLINE=20036019; PubMed=10571241;
RA Hiesberger T., Trommsdorff M., Howell B.W., Goffinet A.M., Mumby M.C.,
RA Cooper J.A., Herz J.;
RT "Direct binding of Reelin to VLDL receptor and ApoE receptor 2 induces
RT tyrosine phosphorylation of disabled-1 and modulates tau
RT phosphorylation.";
RL Neuron 24:481-489(1999).
CC -|- FUNCTION: Binds VLDL and transports it into cells by endocytosis.
CC In order to be internalized, the receptor-ligand complexes must
CC first cluster into clathrin-coated pits. Binding to Reelin induces
CC tyrosine phosphorylation of Dab1 and modulation of Tau
CC phosphorylation.
CC -|- SUBUNIT: Binds to the extracellular matrix protein Reelin.
CC -|- SUBCELLULAR LOCATION: Type I membrane protein.
CC -|- TISSUE SPECIFICITY: ABUNDANT IN HEART AND MUSCLE; LESS IN KIDNEY,
CC BRAIN, OVARY, TESTIS, LUNG, ADIPOSE TISSUE.
CC -|- SIMILARITY: CONTAINS 8 LDL-RECEPTOR CLASS A DOMAINS.
CC -|- SIMILARITY: CONTAINS 6 LDL-RECEPTOR CLASS B DOMAINS.
CC -|- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----

```

```

DR EMBL: L33417; AAC37668.1; -.
DR DR EMBL: U06670; AAA59384.1; -.
DR DR EMBL: S73732; AAB32228.2; -.
DR DR HSPD; P01130; 1AJJ.
DR DR MGD; MGI:98935; Vldlr.
DR DR InterPro: IPR000152; Asx_hydroxyl.
DR DR InterPro: IPR000561; EGF-like.
DR DR InterPro: IPR001881; EGF_Ca.
DR DR InterPro: IPR002172; LDL_recept_A.
DR DR InterPro: IPR000033; Ldl_receptor_rep.
DR DR Pfam; PF00057; ldl_recept_a; 8.
DR DR Pfam; PF00058; ldl_recept_b; 5.
DR DR PRINTS; PR00281; LDLRECEPTOR.
DR DR SMART; SM00179; EGF_CA; 2.
DR DR SMART; SM00001; EGF_like; 1.
DR DR SMART; SM00192; LDLa; 8.
DR DR SMART; SM00135; LY; 5.
DR DR PROSITE; PS00010; ASX_HYDROXYL; 2.
DR DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR DR PROSITE; PS01186; EGF_2; 3.
DR DR PROSITE; PS01187; EGF_CA; 1.
DR DR PROSITE; PS01209; LDLRA_1; 8.
DR DR PROSITE; PS00068; LDLRA_2; 8.
DR DR Glycoprotein; VLDL; Cholesterol metabolism; Lipid transport;
KW Endocytosis; Coated pits; Transmembrane; Receptor; Signal;
KW EGF-like domain; Repeat.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 873 VERY LOW-DENSITY LIPOPROTEIN RECEPTOR.
FT DOMAIN 28 797 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 798 819 POTENTIAL.
FT DOMAIN 820 873 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 31 69 LDL-RECEPTOR CLASS A 1.
FT DOMAIN 70 110 LDL-RECEPTOR CLASS A 2.
FT DOMAIN 111 151 LDL-RECEPTOR CLASS A 3.
FT DOMAIN 152 190 LDL-RECEPTOR CLASS A 4.
FT DOMAIN 191 231 LDL-RECEPTOR CLASS A 5.
FT DOMAIN 232 275 LDL-RECEPTOR CLASS A 6.
FT DOMAIN 276 314 LDL-RECEPTOR CLASS A 7.
FT DOMAIN 315 355 LDL-RECEPTOR CLASS A 8.
FT DOMAIN 356 391 EGF-LIKE 1.
FT DOMAIN 392 431 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT REPEAT 432 480 LDL-RECEPTOR CLASS B 1.
FT REPEAT 481 524 LDL-RECEPTOR CLASS B 2.
FT REPEAT 525 567 LDL-RECEPTOR CLASS B 3.
FT REPEAT 568 611 LDL-RECEPTOR CLASS B 4.
FT REPEAT 612 654 LDL-RECEPTOR CLASS B 5.
FT REPEAT 655 696 LDL-RECEPTOR CLASS B 6.
FT DOMAIN 702 750 EGF-LIKE 3.
FT DOMAIN 751 790 CLUSTERED O-LINKED OLIGOSACCHARIDES.
FT SITE 832 837 ENDOCYTOSIS SIGNAL (POTENTIAL).
FT DISULFID 33 45 BY SIMILARITY.
FT DISULFID 40 58 BY SIMILARITY.
FT DISULFID 52 67 BY SIMILARITY.
FT DISULFID 72 84 BY SIMILARITY.
FT DISULFID 79 97 BY SIMILARITY.
FT DISULFID 91 108 BY SIMILARITY.
FT DISULFID 113 127 BY SIMILARITY.
FT DISULFID 120 140 BY SIMILARITY.
FT DISULFID 134 149 BY SIMILARITY.
FT DISULFID 154 166 BY SIMILARITY.
FT DISULFID 161 179 BY SIMILARITY.
FT DISULFID 173 188 BY SIMILARITY.
FT DISULFID 193 205 BY SIMILARITY.
FT DISULFID 200 218 BY SIMILARITY.
FT DISULFID 212 229 BY SIMILARITY.
FT DISULFID 239 251 BY SIMILARITY.
FT DISULFID 246 264 BY SIMILARITY.
FT DISULFID 258 273 BY SIMILARITY.
FT DISULFID 278 290 BY SIMILARITY.
FT DISULFID 285 303 BY SIMILARITY.
FT DISULFID 297 312 BY SIMILARITY.
FT DISULFID 318 331 BY SIMILARITY.
FT DISULFIQ 326 344 BY SIMILARITY.

FT DISULFID 338 355 BY SIMILARITY.
FT DISULFID 360 371 BY SIMILARITY.
FT DISULFID 367 380 BY SIMILARITY.
FT DISULFID 382 394 BY SIMILARITY.
FT DISULFID 400 410 BY SIMILARITY.
FT DISULFID 406 419 BY SIMILARITY.
FT DISULFID 421 434 BY SIMILARITY.
FT DISULFID 706 719 BY SIMILARITY.
FT DISULFID 715 734 BY SIMILARITY.
FT DISULFID 736 749 BY SIMILARITY.
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 765 765 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 781 781 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 161 161 C -> G (IN REF. 1).
FT CONFLICT 262 262 P -> L (IN REF. 3).
FT CONFLICT 297 297 C -> S (IN REF. 2).
SQ SEQUENCE 873 AA; 96372 MW; 08F09F93825195CB CRC64;

Query Match 6.7%; Score 89; DB 1; Length 873;
Best Local Similarity 22.5%; Pred. No. 6.8; Indels 54; Gaps 11;
Matches 48; Conservative 29; Mismatches 82;

QY 44 HVPNSVPLKEVLMKQKQKVAELENSEF-----RAPSFKNRYVLDTVSG-SUTIYNLAFSS 98
Db 622 HMLSSVDL-----NGQDRIRVLKSLFLAHLALTFEDRVY--WIDGENEAVYKANKF 673
QY 99 DEDEVEMESPNITDTMKFFLYVLESPLSPTLTCA--LTNGSIEVQCMPIEHNHSHRGLIM 156
Db 674 TGSEALTNVNLNDAQDIIVYHELVPQSGKNWCEDDMENGCEVLCUPAPQINDHSP--K 731
QY 157 YSWDCPM-----EOCKRNSTSIYFKMENDLPQKIQCTLSNPLFNWT-----S 198
Db 732 YTCSPNGYNLEENGRCQSTSTPVYSETKDI-----NTTDLILRTSLVPG 778
QY 199 STILTTC-----IPSSGSHRHYALIPPLAVI 226
Db 779 GINVTAVSEVSPPKGTSA-AWAILPLLLLVM 810

RESULT 15
LDVR_RAT STANDARD; PRT; 873 AA.
ID LDVR_RAT AC P98166;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Very low-density lipoprotein receptor precursor (VLDL receptor).
GN VLDLR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95014480; PubMed=7929362;
RA Jokinen E.V., Landschulz K.T., Wyne K.L., Ho Y.K., Frykman P.K.,
RA Hobbs H.H.;
RT "Regulation of the very low density lipoprotein receptor by thyroid
RT hormone in rat skeletal muscle."
RL J. Biol. Chem. 269:26411-26418(1994).
CC -!- FUNCTION: Binds VLDL and transports it into cells by endocytosis.
CC In order to be internalized, the receptor-ligand complexes must
CC first cluster into clathrin-coated pits. Binding to Reelin induces
CC tyrosine phosphorylation of Dab1 and modulation of Tau
CC phosphorylation (By similarity).
CC -!- SUBUNIT: Binds to the extracellular matrix protein Reelin (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: ABUNDANT IN SOLEUS, GASTROCNEMICUS, HEART
CC MUSCLE, PLACENTA, BRAIN, LUNG AND WHITE FAT. LESS IN AORTA, OVARY,
CC KIDNEY, SPLEEN, ADRENAL GLAND, AND THYMUS.
CC -!- SIMILARITY: CONTAINS 8 LDL-RECEPTOR CLASS A DOMAINS.

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: January 28, 2003, 08:38:45 : Search time 52.7638 Seconds
(without alignments)
976.271 Million cell updates/sec

Title: US-09-730-465-2

Sequence: 1 MVAGSDAGRALGVLSVVCLL.....VLYMNGILKCDKRPDRTNSN 250

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1244	93.8	240	4	Q9BRW0 homo sapien
2	647	48.8	134	4	Q14748 homo sapien
3	562	42.4	119	4	Q16393 homo sapien
4	541.5	40.8	253	6	Q28753 ovis sp. lf
5	517.5	39.0	227	6	Q28754 ovis sp. lf
6	453.5	34.2	244	6	Q8SQB6 sus scrofa
7	326.5	24.6	159	6	Q28752 ovis sp. lf
8	197.5	14.9	357	13	Q90ZL5 anas platyr
9	124.5	9.4	329	11	Q92178 mus musculus
10	115.5	8.7	351	4	Q96TE5 homo sapien
11	114.5	8.6	288	6	Q28499 macaca mula
12	112	8.4	288	6	Q77684 macaca neme
13	106.5	8.0	833	5	Q9BP07 halocynthia
14	105.5	8.0	311	11	Q9JLM2 rattus norv
15	105	7.9	285	11	Q9D780 mus musculus
16	105	7.9	285	11	Q8VE93 mus musculus

17	104.5	7.9	536	4	Q96RW7	Q96rw7 homo sapien
18	103	7.8	336	6	Q8WNV8	Q8wnv8 felis silve
19	102.5	7.7	1252	4	Q96DN3	Q96dn3 homo sapien
20	102	7.7	335	13	Q9PWR4	Q9pwr4 gallus gall
21	100.5	7.6	377	10	Q9SKK8	Q9skk8 arabidopsis
22	100.5	7.6	1434	10	Q9AW39	Q9aw39 quillardia
23	99	7.5	234	11	Q61401	Q61401 mus musculus
24	99	7.5	335	13	Q9YGH1	Q9ygh1 gallus gall
25	98.5	7.4	415	11	Q60977	Q60977 mus musculus
26	97.5	7.4	181	13	Q91665	Q91665 xenopus lae
27	97.5	7.4	276	16	Q9ZDS2	Q9zds2 rickettsia
28	97	7.3	197	4	Q9UIT7	Q9uit7 homo sapien
29	97	7.3	241	4	Q9UIB6	Q9uib6 homo sapien
30	97	7.3	272	4	Q9UIB7	Q9uib7 homo sapien
31	97	7.3	280	4	Q95660	Q95660 homo sapien
32	97	7.3	328	4	Q15430	Q15430 homo sapien
33	97	7.3	328	7	Q8WLP1	Q8wlp1 homo sapien
34	97	7.3	335	13	Q9YGV5	Q9ygv5 gallus gall
35	97	7.3	339	4	Q8WWI8	Q8wwi8 homo sapien
36	97	7.3	345	4	Q9UIB8	Q9uib8 homo sapien
37	96.5	7.3	318	13	Q91664	Q91664 xenopus lae
38	96.5	7.3	323	6	Q9BDM2	Q9bdm2 cercopithe
39	95	7.2	2489	3	Q06116	Q06116 saccharomyc
40	94.5	7.1	204	11	Q9JLM3	Q9jlm3 rattus norv
41	94.5	7.1	261	11	Q9D6A9	Q9d6a9 mus musculus
42	94.5	7.1	332	13	Q9IB08	Q9ib08 spherooides
43	94.5	7.1	863	15	Q8Q7G5	Q8q7g5 human immun
44	94	7.1	408	11	Q91WP1	Q91wpl mus musculus
45	93.5	7.1	333	13	Q9IB04	Q9ib04 spherooides

ALIGNMENTS

RESULT 1

Q9BRW0
ID Q9BRW0 PRELIMINARY; PRT; 240 AA.
AC Q9BRW0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Similar to CD58 antigen, (lymphocyte function-associated antigen 3).
DE 3).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRIN;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC005930; AAH05930.1; -.
DR HSSP; P19256; ICCZ.
DR InterPro; IPR003599; Ig.
DR SMART; SM00409; IG; 1.
SQ SEQUENCE 240 AA; 27044 MW; 431E44EFEDDF80B2 CRC64;

Query Match 93.8%; Score 1244; DB 4; Length 240;
Best Local Similarity 99.2%; Pred. No. 1.7e-110;
Matches 235; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVAGSDAGRALGVLSVVCLLHCFGFISCFSSQOIYGVVYGNVTFHVPNSVPLKVLWKQK 60

Db 1 MVAGSDAGRALGVLSVVCLLHCFGFISCFSSQOIYGVVYGNVTFHVPNSVPLKVLWKQK 60

Qy 61 DKVALENSERAFSSKKNVYLDVTGSLTYNTSSDEDEYEMESPNTDTMKFFLYV 120

Db 61 DKVALENSERAFSSKKNVYLDVTGSLTYNTSSDEDEYEMESPNTDTMKFFLYV 120

Qy 121 LESLSPPTLTCAITNGSIEVQCQIPEYNSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180

Db 121 LESLSPPTLTCAITNGSIEVQCQIPEYNSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180

```
QY 181 LPKIQCTLSNPLNFTTSSIIITTCIPSSGSHRHYALIPILAVITTCIVLYMNGI 237
Db 181 LPKIQCTLSNPLNFTTSSIIITTCIPSSGSHRHYALIPILAVITTCIVLYMNGM 237

RESULT 2
Q14748 ID Q14748 PRELIMINARY; PRT; 134 AA.
AC Q14748;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DE LFA-3(delta D2) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA kakutani T.;
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; D28586; BAA05922.1; -.
DR HSSP; P19256; ICCZ.
DR InterPro; IPR003599; Ig.
DR SMART; SM00409; Ig; 1.
FT NON_TER 1
SQ SEQUENCE 134 AA; 15452 MW; 79BEF0A4EEB4E59B CRC64;

Query Match 48.8%; Score 647; DB 4; Length 134;
Best Local Similarity 60.4%; Pred. No. 6.2e-54;
Matches 134; Conservative 0; Mismatches 0; Indels 88; Gaps 1;

QY 29 FSQIQYGVYGVNTHVPSNVPLKEVLWKQKDKVAELENSEFRASFKNRVYLDTVSG 88
Db 1 FSQIQYGVYGVNTHVPSNVPLKEVLWKQKDKVAELENSEFRASFKNRVYLDTVSG 60

QY 89 SLTIYNTSSDEDEYEMESPNITDTMKFFLYLVLESPLTLCALTNGSIEVQCMPEHY 148
Db 61 SLTIYNTSSDEDEYEMESPNITDTMKFFLYLV----- 93

QY 149 NSHRGLIMYSWDCPMEQCKRNSTSIYFKMENDLPQKIQCTLSNPLNFTTSSIIITTCIPS 208
Db 94 ----- 93

QY 209 SGHSRHYALIPILAVITTCIVLYMNGILKCDKPKDRTNSN 250
Db 94 -GHSRHYALIPILAVITTCIVLYMNGILKCDKPKDRTNSN 134

RESULT 3
Q16393 ID Q16393 PRELIMINARY; PRT; 119 AA.
AC Q16393;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DE LFA-3 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=95398636; PubMed=7545392;
RA Kirby A.C., Hill V., Olsen I., Porter S.R.;
RT "LFA-3 delta D2: a novel in vivo isoform of lymphocyte function-
associated antigen 3.";
RL Biochem. Biophys. Res. Commun. 214:200-205(1995).
DR EMBL; S79616; ABA0000.2; -.
DR HSSP; P19256; ICCZ.
FT NON_TER 1
NON_TER 119 119
```

```
SQ SEQUENCE 119 AA; 13732 MW; 4096AC6F45181505 CRC64;

Query Match 42.4%; Score 562; DB 4; Length 119;
Best Local Similarity 57.0%; Pred. No. 6.7e-46;
Matches 118; Conservative 0; Mismatches 1; Indels 88; Gaps 1;

QY 37 VYGNVTHVPSNVPLKEVLWKQKDKVAELENSEFRASFKNRVYLDTVSGSLTIYNLT 96
Db 1 VYGNVTHVPSNVPLKEVLWKQKDKVAELENSEFRASFKNRVYLDTVSGSLTIYNLT 60

QY 97 SSDEDEYEMESPNITDTMKFFLYLVLESPLTLCALTNGSIEVQCMPEHYNSHRGLIM 156
Db 61 SSDEDEYEMESPNITDTMKFFLYLV----- 85

QY 157 YSWDCPMEQCKRNSTSIYFKMENDLPQKIQCTLSNPLNFTTSSIIITTCIPSSGSHRHY 216
Db 86 -----GHSRHY 92

QY 217 ALIPIPLAVITTCIVLYMNGILKCDK 243
Db 93 ALIPIPLAVITTCIVLYMNGILKCDK 119

RESULT 4
Q28753 ID Q28753 PRELIMINARY; PRT; 253 AA.
AC Q28753;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE LFA-3 precursor.
OS Ovis sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9939;
RN [1]
RP SEQUENCE FROM N.A.
RA kakutani T.;
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; D28584; BAA05920.1; -.
DR HSSP; P19256; ICCZ.
DR InterPro; IPR001993; Mitoch_carrier.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_1.
KW Signal.
FT SIGNAL 1 28 POTENTIAL.
SQ SEQUENCE 253 AA; 27715 MW; 4A4E97A38518B290 CRC64;

Query Match 40.8%; Score 541.5; DB 6; Length 253;
Best Local Similarity 48.8%; Pred. No. 1.5e-43;
Matches 121; Conservative 34; Mismatches 78; Indels 15; Gaps 8;

QY 1 MVAGSDAGRALGVLSVVCLLHCGFISCFSQIYGVVGVNTHVPSNVPLKEVLWKQK 60
Db 1 WAAGSAPGCALGALGLVCLFLKLDIFSCVSDIYGAMGNVTFYVSESQPFTEIMWKKG 60

QY 61 DKVAEL-ENSEFRASFKNRVYLDTVSGSLTIYNLTSSDEDEYEMESPNITDTMKFFLY 119
Db 61 DKVVEWDQTSGLAEAFOSFKNRVHLDIVSGNLTITGLTKLDEYIEESPVSVKKSQFH 120

QY 120 VLESPLTTCALT-NGSIEVQCMPE---HYNSHRGLIMYSWDCPME-QCKRNST-- 172
Db 121 VIEPPTPSASCFTEGGNITLTCSIPEGDPKELDDSDILRYLWECPTTQCHRSISSE 180

QY 173 IYFKMENDLPQKIQCTLSNPLNFTTSSIIITTCIPSSGSHRHYALIPILAVITTCIVL 232
Db 181 AFVSAESDLSQNVQICVSNPLFRTSASVSLTCLPED-YAHRHYVLFATLPAVI--CGLL 237

QY 233 YMNGILK 240
Db 236 F-----LKC 241
```

RESULT 5

Q28754 ID Q28754 PRELIMINARY; PRT; 227 AA.
 AC Q28754; 1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE LFA-3(delta TM) precursor.
 OS Ovis sp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9939;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kakutani T.;
 RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
 DR EMBL; D28585; BAA05921.1; -.
 DR HSSP; P19256; ICCZ.
 KW Signal.
 FT SIGNAL 29 227 POTENTIAL.
 FT CHAIN 1 28 LFA-3(Delta TM).
 SQ SEQUENCE 227 AA; 24760 MW; 1521A800B49E9B55 CRC64;

Query Match 39.0%; Score 517.5; DB 6; Length 227;
 Best Local Similarity 48.9%; Pred. No. 2.5e-41;
 Matches 110; Conservative 34; Mismatches 72; Indels 9; Gaps 6;

QY 1 MVAGSDAGRALGVLSVVCLLHCFGIFSCFSQIYGVYGVNTHFVPSNVPLKEVLWKQK 60
 Db 1 MAAGSAPGCALGALGVLCFLKDFISCVSQDIYGAMNGNVTFYVSESQPTETIMWKKG 60
 QY 61 DKVAEL-ENSEFRASFKNRVYLDTVSGSLTIYNLTSSDEDEYEMESPNITDTMKFFLY 119
 Db 61 DKVVEWDQTSGLAEAFQSFKNRVHLDIVSGNLITGLTKLDEDVYEIESPSVKKSQFHLR 120
 QY 120 VLESPLPTTLCALT-NGSIEVOCMIPF---HYNHSHRGLIMYSWDCPME--OCKRNSTSIYFK 176
 Db 121 VIEPPPTASCFLETGGNITLTCISPEGDPKELDDSDIRYLWECPTTIOCHRGSISS 180
 QY 173 IYFMENDLPKIQCTLSNPLNTSSIIITTCIPSSGSHRYALIPLAVITTCIVL 232
 Db 181 AFVSAESDLSNVQCIVSNPLFTSASVSLTCLPED-VARHRS 224

RESULT 6

Q8SQB6 ID Q8SQB6 PRELIMINARY; PRT; 244 AA.
 AC Q8SQB6;
 DT 01-JUN-2002 (TReMBLrel. 21, Created)
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE CD58 antigen.
 GN CD58.
 OS Sus scrofa (pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Brossay A., Hube F., Bardos P., Watier H.;
 RT "Characterization of the porcine CD58 antigen mRNA sequence."
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF469666; AAL76278.1; -.
 SQ SEQUENCE 244 AA; 27301 MW; 164B81ACC7A512D2 CRC64;

Query Match 34.2%; Score 453.5; DB 6; Length 244;
 Best Local Similarity 40.7%; Pred. No. 3.4e-35;
 Matches 96; Conservative 42; Mismatches 93; Indels 5; Gaps 3;

QY 1 MVAGSDAGRALGVLSVVCLLHCFGIFSCFSQIYGVYGVNTHFVPSNVPLKEVLWKQK 60
 Db 1 MASGRALRAVAPDVFILLMLDFIHCDSKVIYGAHLKHNVTURASSQTQETIIMWKKG 60

QY 61 DKVAELNS-EFRAFSSFKNRVYLDTVSGSLTIYNLTSSDEDEYEMESPNITDTMKFFLY 119
 Db 61 DKAVEWEQYNVTAFFPPFVDRVHLNTGSDGLTIYNLAFEDGDIYESPSVKNSSKFIILR 120
 QY 120 VLESPLPTTLCALTNGSIEVOCMIPHEYNHSHRGLIMYSWDCPME---OCKRNSTSIYFK 176
 Db 121 VVEPLPEPELYCESTEIGNISVRCLILAAASHRIDLIQYWNCPPTVSCQAGLGPSEMYIT 180
 QY 177 MENLDPKIQCTLSNPLNTSSIIITTCIPSSGSHRYALIPLAVITTCIVL 232
 Db 181 KESDFSQDVQCIISNPLFKSSASLUSCAPID-NTRHRIVLVAVLPAVIALVLL 235

RESULT 7

Q28752 ID Q28752 PRELIMINARY; PRT; 159 AA.
 AC Q28752;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE LFA-3(delta D2) precursor.
 OS Ovis sp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9939;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kakutani T.;
 RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
 DR EMBL; D28583; BAA05919.1; -.
 DR HSSP; P19256; ICCZ.
 KW Signal.
 FT SIGNAL 1 28 POTENTIAL.
 SQ SEQUENCE 159 AA; 17645 MW; 592A97046ECD985E CRC64;

Query Match 24.6%; Score 326.5; DB 6; Length 159;
 Best Local Similarity 45.4%; Pred. No. 2.6e-23;
 Matches 74; Conservative 22; Mismatches 52; Indels 15; Gaps 4;

QY 1 MVAGSDAGRALGVLSVVCLLHCFGIFSCFSQIYGVYGVNTHFVPSNVPLKEVLWKQK 60
 Db 1 MAAGSAPGCALGALGVLCFLKDFISCVSQDIYGAMNGNVTFYVSESQPTETIMWKKG 60
 QY 61 DKVAEL-ENSEFRASFKNRVYLDTVSGSLTIYNLTSSDEDEYEMESPNITDTMKFFL- 118
 Db 61 DKVVEWDQTSGLAEAFQSFKNRVHLDIVSGNLITGLTKLDEDVYEIESPSVKKSQFHLR 120
 QY 119 -----VLESPLPTTLCALTNGSIEVOCMIPHEYNHSHRG 153
 Db 121 VIDYARHRYVLFAT-LPAVIC----GLLFLKCFGLGRRSQRNSG 158

RESULT 8

Q90ZL5 ID Q90ZL5 PRELIMINARY; PRT; 357 AA.
 AC Q90ZL5;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE CD58 antigen.
 OS Anas platyrhynchos (Domestic duck).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.
 OX NCBI_TaxID=8839;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-WHITE PEKIN;
 RA Chan S.W.S., Middleton D.L., Lundqvist M., Warr G.W., Higgins D.A.;
 RT "Anas platyrhynchos T cell antigens."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY032731; AAK51607.1; -.

DR	InterPro: IPR003006; Ig_MHC.	
DR	Pfam: PF00047; Ig_2.	
SQ	SEQUENCE 357 AA; 39930 MW; 274161DF93868C08 CRC64;	
	Query Match 14.9%; Score 197.5; DB 13; Length 357;	
	Best Local Similarity 28.0%; Pred. No. 1.3e-10;	
	Matches 61; Conservative 41; Mismatches 101; Indels 15; Gaps 6;	
QY	26 ISCFSDIYGVVYGVNTHFVPSNVPLKEVLWKKOKDKVAELE-NSEFRAFSFKNRVYLD 84	
DB	22 IYC-EKVPGLIGENFTFPVKADNKIDVITWKDKDKVAEWAQSEPTYFNSLRSLLN 80	
QY	85 TVSGSTIYNTLTSDEDEYEM---SPNITDTMKFFLVLESPLSPPLTLCALNGSIEVO 141	
DB	81 KESGNITIFKLENSDGMVHLERFSRTENGVTFLNLTVDLPPSEPOINGSLSDNNMLT 140	
QY	142 CMIPESHYNSHRLIMYSWDCPMEQCKRNSTSIYFKWENDLPKIQCTLSNPLNFTSSIT 201	
DB	141 CEA-----KFQPLTYVWKITGRIFTGPKVLIPKENVDTGKATCFVYFKMTKSSEIT 195	
QY	202 LTTCTIPS-SGHSRHRVYALIPILAVITTCIV-----LYM 234	
DB	196 LDQCSPTYQGYSSPKRSRAGLIITAVAFVFGAGLIYM 233	
	RESULT 9	
ID	Q92178 PRELIMINARY; PRT; 329 AA.	
AC	Q92178;	
DT	01-MAY-1999 (Tremblrel. 10, Created)	
DT	01-MAY-1999 (Tremblrel. 10, Last sequence update)	
DT	01-DEC-2001 (Tremblrel. 19, Last annotation update)	
DE	CD84 leukocyte antigen.	
GN	CD84 OR CD84.	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
OX	NCBI_TaxID=10090;	
RN	[1]	
RC	SEQUENCE FROM N.A.	
RP	TISSUE=PERITONEUM;	
RX	MEDLINE=99180614; PubMed=10079287;	
RA	de la Fuente M.A., Tovar V., Pizcueta P., Nadal M., Bosch J.,	
RA	Engel P.;	
RT	"Molecular cloning, characterization, and chromosomal localization of	
RT	the mouse homologue of CD84, a member of the CD2 family of cell	
RT	surface molecules.";	
RL	Immunogenetics 49:249-255(1999).	
DR	EMBL: AF043445; AAD02273.1; -.	
DR	MGI:1336885; CD84.	
DR	InterPro: IPR003600; Ig_Like.	
DR	SMART: SM00410; Ig_Like.1.	
SQ	SEQUENCE 329 AA; 37345 MW; 43BB1AA5AF1989E0 CRC64;	
	Query Match 9.4%; Score 124.5; DB 11; Length 329;	
	Best Local Similarity 25.1%; Pred. No. 0.0011;	
	Matches 56; Conservative 43; Mismatches 93; Indels 31; Gaps 13;	
QY	35 GVVYGVNTHFVPSNVPLK--EVLWKKOKDKVAELENSSEFRA-----FSSFKNRVYLDTVS 87	
DB	31 GILGESVFLNIQEPKKIDNIATSQ-SSVAFIPGVNKAEVITOGTYKGRIFDIQK 89	
QY	88 GSITINYNTLTSDEDEYEM---ESPNTDTMKFFLVLESPLSPPLTLCALNG-----SIE 139	
DB	90 YDLVTRDLRMEADAGTKADINEEETITKIYLIHYRLKTPKTIQSLSSLNNTCNIT 149	
QY	140 VOCMPEHYNSHRLIMYSWDCPMEQCKRNSTSIYFKWENDLPKI--OCTLSNPLNFT 197	
DB	150 LACSVEKEKD-----VYIWSN-PFGE-KSNVLQI---VHSPMDQKLIYTTCTAQNPSNSS 200	
QY	198 SSI-ILTTCTIPS-SGHSRHRVYALIPILAVITTCIVLYMNGIL 238	
DB	DSVTVOQPTDTPSPFHPRH--AVLPGGLAVFLFLILIPLAFL 241	
	RESULT 10	
ID	Q96TES PRELIMINARY; PRT; 351 AA.	
AC	Q96TES;	
DT	01-DEC-2001 (Tremblrel. 19, Created)	
DT	01-DEC-2001 (Tremblrel. 19, Last sequence update)	
DT	01-JUN-2002 (Tremblrel. 21, Last annotation update)	
DE	DJ655N15.1 (CD2 antigen (p50), sheep red blood cell	
DE	receptor).	
GN	CD2.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	Hall R.;	
RL	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.	
DR	EMBL: ALI35798; CAC14840.1; -.	
SQ	SEQUENCE 351 AA; 39448 MW; A03D853C3B618917 CRC64;	
	Query Match 8.7%; Score 115.5; DB 4; Length 351;	
	Best Local Similarity 23.0%; Pred. No. 0.0084;	
	Matches 59; Conservative 42; Mismatches 76; Indels 79; Gaps 17;	
QY	32 QIYGVYGVNTHFVPS---NVPLKEVLWKKOKD--KVAEL--ENSEFR---AFSFKNRV 81	
DB	32 ETWGALGQDINLDIPSFQMSDDIDDIKWETSDKKKIAQFRKEKTFREKDYKLFKN-- 89	
QY	82 YLDTVSGSLIYNTLTSDEDEYEM-----ESPNTDTMKFFLVLESPLSP----- 127	
DB	90 -----GTLKIKHLKTDQDIYKVSIVDTKGKNVLEKI-FDLKIQERVSKPKISWTCINT 142	
QY	128 TLTCALNGSIEVQCMPE-----HNSHRGLIMYSWDCPMEQCKRNSTSIYFKMEN 179	
DB	143 TLTCVWNGT-----DPELNLYQDGKHLKLSQRVITHKW-----FTS----- 179	
QY	180 DLPQIQCTLSNPLNFTTSSIIITCIPSSGHSRHRVYALIPI-----PLAVITTCIVLYM 234	
DB	180 -LSAKFKCTAGNKV-SKESSEVPVSC-PEKG--LDIYLIIGICGGGSLLMVFPVALLVFI 234	
QY	235 NGILKCDRKPDRTNSN 250	
DB	235 T-----RRKKQSRRN 245	
	RESULT 11	
ID	Q28499 PRELIMINARY; PRT; 288 AA.	
AC	Q28499;	
DT	01-NOV-1996 (Tremblrel. 01, Created)	
DT	01-NOV-1996 (Tremblrel. 01, Last sequence update)	
DT	01-JUN-2002 (Tremblrel. 21, Last annotation update)	
DE	B7 protein (CD80 protein precursor).	
GN	B7 OR N939.	
OS	Macaca mulatta (Rhesus macaque).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;	
OC	Cercopitheciinae; Macaca.	
OX	NCBI_TaxID=9544;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=BLOOD;	
RX	MEDLINE=96003435; PubMed=7561102;	
RA	Villinger F., Brar S.S., Wayne A., Chikkala N., Ansari A.A.;	
RT	"Comparative sequence analysis of cytokine genes from human and	
RT	nonhuman primates.";	
RL	J. Immunol. 155:3946-3954(1995).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RA	Villinger F., Bostik P., Mayne A.E., King C.L., Genain C.P.,	


```

RESULT 14
Q9JLM2 ID Q9JLM2 PRELIMINARY; PRT; 311 AA.
AC Q9JLM2;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE NK cell receptor 2B4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20260988; PubMed=10803843;
RA Kumarasan P.R., Stepp S.E., Bennett M., Kumar V., Mathew P.A.;
RT "Molecular cloning of transmembrane and soluble forms of a novel rat
RL natural killer cell receptor related to 2B4.";
RL Immunogenetics 51:306-313(2000).
DR EMBL; AF156989; AAF71162.1; -.
KW Receptor.
SQ SEQUENCE 311 AA; 35300 MW; 5A527D2D2D2565772 CRC64;

Query Match 8.0%; Score 105.5; DB 11; Length 311;
Best Local Similarity 21.6%; Pred. No. 0.065;
Matches 59; Conservative 56; Mismatches 99; Indels 59; Gaps 16;

Qy 13 VLSVCLLHCFGIFSC--FSQIYGVVYGNVTFHVPSPNVLKEVL--WKK-----QKDKV 63
Db 7 LLSLFLLLRAHQGDQCADSSEELGVSGKPVRLR-PSNIQAKHVSIEWKKTGHQQTQI 65

Qy 64 AELESEPRAFSSFNRYV-LDVTSGSLTYNLTSSDEDEYEMESPN---ITDTMKFFLY 119
Db 66 VTWDTDNKFNWCCSDIYGFESENFALSIAKLNDSGHYLLLEITNQRGIVCTKNFQML 125

Qy 120 VLESPLSPITLCALT---NGS--IEVOCIPE-----HYNHSHRGLIMYS-----WDCPM 163
Db 126 IFDPVETPLHTVQGTFLWANGTCLSLSCFVPKDDNVSYALYRGSMILSNQRYGTHW---- 181

Qy 164 EQCKRNSYIFKMNELPKQIOCTLSNPLF---NTTSSIIITLCIPSSGHSRHRVALIP 220
Db 182 -ENRTDASSLH-----TYTCNVSNKASWANHTLTPSQSCQSVP-----SKFNY---- 223

Qy 221 IPLAVITTCIVLYMNGILKC-----DRKPDRTNS 249
Db 224 LPFMVSGILVKKFFHGAIIDCFVNNRRKRKQSQS 256

RESULT 15
Q9D780 ID Q9D780 PRELIMINARY; PRT; 285 AA.
AC Q9D780;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE 2310026I04Rik protein.
GN CD2F10 OR 2310026I04RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=TONGUE;
RA MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamakawa I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Futschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

```

```

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weicz C., Whittaker C., Wilming L.,
RA Wuyshaw-Boriss A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK009505; BAB26328.1; -.
DR MGD; MGI:1923692; Cg2f10.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003600; Ig_Like.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00409; IG; 1.
DR SMART; SM00410; IG_Like; 1.
SQ SEQUENCE 285 AA; 31764 MW; 33BDB4A633A5C1B3 CRC64;

```

Query Match 7.9%; Score 105; DB 11; Length 285;

Best Local Similarity 22.1%; Pred. No. 0.065;
Matches 45; Conservative 29; Mismatches 74; Indels 56; Gaps 7;

```

Qy 31 QQIYGVVYG--NVTFFHVPSPNVLKEVLW-----KKOKDKVAELENSEFRAPSS 76
Db 26 EEVGVQLQESINLSLEIPSNEEIKHIDWLFQNNIAIVKPKKGQPAVIMAVDPRYR---- 81

Qy 77 FKNRVYLDVTSGSLTYNLTSSDEDEY---EMESPNITDITMKFFLYLVELESPLTLTCA 132
Db 82 --GRVSISSSYSLHLSNLTWEDSCLYNAQVNLKTSSEHTKSYHLRYRLSKPHIT-- 137

Qy 133 LTNGSIEVOCMPEHYHNSHRGLIMYSWDCPMEQCKRNSYIFKMN----- 179
Db 138 -VNSNI-----SEEGVCNISLTCSIERAGMDVTYIMLSSQDSTNTSHESGVLSLT 185

Qy 180 -----DLQPKIOCTLSNPLFNFTS 198
Db 186 SWRPGDKAPSYTCRVSNPISNIS 209

```

Search completed: January 28, 2003, 08:44:16
Job time : 56.7638 secs

t

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 08:38:45 ; Search time 18.0184 seconds
(without alignments)
1851.363 Million cell updates/sec

Title: US-09-730-465-8

Perfect score: 1856

Sequence: 1 MVAGSDAGRALGVLSVCLL.....MHEALHNHYTKSLSPGK 347

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR.73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1233	66.4	255	4 S31866	Ig gamma-1 chain C
2	1233	66.4	330	1 GHU	Ig gamma-1 chain C
3	1228.5	66.2	374	2 S69339	Ig heavy chain V r
4	1180	63.6	234	2 P70207	Ig gamma chain C r
5	1146	61.7	377	2 A23511	Ig gamma-3 chain C
6	1145.5	61.7	326	1 G2HU	Ig gamma-2 chain C
7	1144	61.6	377	2 A60764	Ig gamma-3 chain C
8	1135	61.2	327	1 G4HU	Ig gamma-4 chain C
9	1121	60.4	289	1 G3HUI	Ig gamma-3 heavy C
10	926.5	49.9	323	1 GHPB	Ig gamma chain C r
11	913.5	49.2	328	2 I47160	Ig gamma 2a chain
12	909.5	49.0	328	2 I47159	Ig gamma 2a chain
13	906.5	48.8	277	2 I47162	Ig gamma 4 chain c
14	894.5	48.2	328	2 I47158	Ig gamma 1 chain c
15	892	48.1	329	1 G2GP	Ig gamma-2 chain C
16	886.5	47.8	328	2 I47161	Ig gamma 3 chain c
17	857.5	46.2	470	2 S22080	Ig heavy chain pre
18	846	45.6	308	2 C30554	Ig heavy chain C r
19	846	45.6	472	2 S31459	Ig gamma-1 chain -
20	845.5	45.6	329	1 G3MSC	Ig gamma-1 chain C
21	838	45.2	333	2 PS0018	Ig gamma-2b chain
22	834.5	45.0	398	1 G3MSM	Ig gamma-3 chain C
23	829.5	44.7	444	2 PC4436	monoclonal antibod
24	819	44.1	326	2 PS0017	Ig gamma-1 chain C
25	818.5	44.1	324	1 G1MS	Ig gamma-1 chain C
26	813.5	43.8	393	1 G1MSM	Ig gamma-1 chain C
27	812	43.8	329	2 S00847	Ig gamma-2c chain
28	811.5	43.7	330	1 G2MSA	Ig gamma-2a chain
29	811.5	43.7	469	2 S37483	Ig gamma-2a chain

ALIGNMENTS

RESULT 1

S31866

Ig gamma-1 chain C region - synthetic

C:Species: synthetic

A:Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli

C:Date: 06-Jan-1995 #sequence_revision 17-Mar-1997 #text_change 19-May-2000

C:Accession: S31866

R:Filpula, D.

submitted to the EMBL Data Library, February 1993

A:Description: Screening method for protein-protein interactions of cloned gene produc

A:Reference number: S31866

A:Accession: S31866

A:Molecule type: mRNA

A:Residues: 1-255 <PI>

A:Cross-references: EMBL:X70421; NID:g33068; PIDN:CAA9866.1; PID:g33069

C:Keywords: immunoglobulin

F:1-22/Region: Escherichia coli outer membrane protein A precursor

F:23-255/Region: human Ig gamma-1 chain C region

Query Match

Best Local Similarity 66.4%; Score 1233; DB 4; Length 255;

Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 DKHTTCCPCAPPELLGGSPVFLPPKPKDTLMTISRTPEVTCVVDVSHEDPEVKFNMYD 180

Db 29 DKHTTCCPCAPPELLGGSPVFLPPKPKDTLMTISRTPEVTCVVDVSHEDPEVKFNMYD 88

Qy 181 GVEVHNAKTPREEQYNSTYRVSVLTVLHODWLNKGYKCKVSNKALPAPIEKTISKAK 240

Db 89 GVEVHNAKTPREEQYNSTYRVSVLTVLHODWLNKGYKCKVSNKALPAPIEKTISKAK 148

Qy 241 GQPREPQVYTLPPSDELTKNQVSLTCLVKGFYPSDIAVEMESNCGPENNYKTTTPPVLD 300

Db 149 GQPREPQVYTLPPSDELTKNQVSLTCLVKGFYPSDIAVEMESNCGPENNYKTTTPPVLD 208

Qy 301 DGSFFLYSKLTVDKSRWQQGNVSCVMHEALNHYTKQSLSPGK 347

Db 209 DGSFFLYSKLTVDKSRWQQGNVSCVMHEALNHYTKQSLSPGK 255

RESULT 2

GHU

Ig gamma-1 chain C region - human

C:Species: Homo sapiens (man)

C:Date: 31-Jan-1981 #sequence_revision 18-Aug-1982 #text_change 16-Jul-1999

C:Accession: A93433; S36861; S33887; B90563; A90564; B91668; A91723; A02146

R:Ellison, J.W.; Berson, B.J.; Hood, L.E.

Nucleic Acids Res. 10, 4071-4079, 1982

A:Title: The nucleotide sequence of a human immunoglobulin C-gamma1 gene.

A:Reference number: A93433

A:Accession: A93433

A:Molecule type: DNA

A:Residues: 1-330 <ELL>
A:Cross-references: EMBL:217370
A:Note: This sequence has the G1m(17) allotypic marker, 97-Lys, and the G1m(1) markers, R:Harris, L.J.
submitted to the EMBL Data Library, October 1992
A:Reference number: S33904
A:Accession: S36861
A:Molecule type: DNA
A:Residues: 2-330 <HAB>
A:Cross-references: EMBL:217370
R:Takahashi, N.; Ueda, S.; Obara, M.; Nikaido, T.; Nakai, S.; Honjo, T.
Cell 29, 671-679, 1982
A:Title: Structure of human immunoglobulin gamma genes: implications for evolution of a A:Reference number: S33887; MUID:83001943; PMID:6811139
A:Accession: S33887
A:Molecule type: DNA
A:Residues: 88-113;235-330 <TAK>
A:Cross-references: EMBL:217370
R:Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman, Biochemistry 9, 3161-3170, 1970
A:Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid sequence A:Reference number: A90563; MUID:71064024; PMID:5489771
A:Contents: myeloma protein Eu
A:Accession: B90563
A:Molecule type: protein
A:Residues: 1-96, 'R', 98-135 <CUN>
A:Note: this sequence has the G1m(3) marker, 97-Arg
R:Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.
Biochemistry 9, 3171-3181, 1970
A:Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid sequence A:Reference number: A90564; MUID:71064025; PMID:5530842
A:Contents: Eu
A:Accession: A90564
A:Molecule type: protein
A:Residues: 136-154, 'O', 156-165, 'O', 167-176, 'O', 178-194, 'N', 196-197, 'D', 199-238, 'E', 240, A:Note: this sequence has the G1m(non-1) markers, 239-Glu and 241-Met
R:Ponstingl, H.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976
A:Title: Die Primaerstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein Nie). igen Primaerstruktur.
A:Reference number: A91668; MUID:77070269; PMID:826475
A:Contents: myeloma protein Nie
A:Accession: B91668
A:Molecule type: protein
A:Residues: 1-34, 'Q', 36-96, 'K', 98-115, 'O', 117-197, 'D', 199-238, 'D', 240, 'L', 242-268, 'E', 27 A:Note: this sequence has the G1m(17) and G1m(1) markers
R:Schmidt, W.E.; Jung, H.D.; Palm, W.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 364, 713-747, 1983
A:Title: Die Primaerstruktur des kristallisierten monoklonalen Immunglobulins IgG1 KO A:Reference number: A91723; MUID:83289131; PMID:6884994
A:Contents: myeloma protein KO; disulfide bonds
A:Accession: A91723
A:Molecule type: protein
A:Residues: 1-96, 'R', 98-197, 'D', 199-238, 'E', 240, 'M', 242-266, 'D', 268-271, 'D', 273-330 <SCH A:Note: this sequence has the G1m(3) and G1m(non-1) markers
R:Gall, W.E.; Edelman, G.M.
Biochemistry 9, 3188-3196, 1970
A:Title: The covalent structure of a human gammaG-immunoglobulin. X. Intrachain disulfid A:Reference number: A90565; MUID:71064027; PMID:4923144
A:Contents: annotation; disulfide bonds
R:Draker, L.; Schwarz, J.; Reichel, W.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 357, 1515-1540, 1976
A:Title: Rule of antibody structure. The primary structure of monoclonal IgG1 immunoglob enonide cleavage products, and the disulfide bridges.
A:Reference number: A91667; MUID:77070267; PMID:1002129
A:Contents: annotation; disulfide bonds
C:Genetics:
A:Gene: GDB:IGHG1
A:Cross-references: GDB:120085; OMIM:147100
A:Map position: 14q32.33-14q32.33
A:Introns: 99/1; 114/1; 224/1
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap

hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:20-85/Domain: immunoglobulin homology <IM1>
F:137-206/Domain: immunoglobulin homology <IM2>
F:243-310/Domain: immunoglobulin homology <IM3>
F:27-83;144-204;250-308/Disulfide bonds: #status experimental
F:103/Disulfide bonds: interchain (to light chain) #status experimental
F:109,112/Disulfide bonds: interchain (to heavy chain) #status experimental
F:180/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 66.4%; Score 1233; DB 1; Length 330;

Best Local Similarity 100.0%; Pred. No. 2e-80;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 DKTHTCPPCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 180
|||||
Db 104 DKTHTCPPCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 163

QY 181 GVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 240
|||||
Db 164 GVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 223

QY 241 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 300
|||||
Db 224 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 283

QY 301 DGSFFLYSKLTVDKSRWQGNVFSCSVHMEALHNHYTQKSLSLSPGK 347
|||||

Db 284 DGSFFLYSKLTVDKSRWQGNVFSCSVHMEALHNHYTQKSLSLSPGK 330

RESULT 3

S69339
Ig heavy chain V region precursor - human
C:Species: Homo sapiens (man)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Dec-2000
C:Accession: S69339; S72664
R:Khamlichi, A.A.; Aucouturier, P.; Preud'homme, J.L.; Cogne, M.
Eur. J. Biochem. 229, 54-60, 1995
A:Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease.
A:Reference number: S69339; MUID:95262687; PMID:7744049
A:Accession: S69339
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-374 <KHA>
A:Cross-references: EMBL:X81695
R:Khamlichi, A.A.
submitted to the EMBL Data Library, September 1994
A:Reference number: S72664
A:Accession: S72664
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-140, 'C', 142-374 <KH2>
A:Cross-references: EMBL:X81695
C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 66.2%; Score 1228.5; DB 2; Length 374;
Best Local Similarity 79.3%; Pred. No. 4.8e-80;
Matches 234; Conservative 10; Mismatches 22; Indels 29; Gaps 1;

QY 82 YLDTVSGSLTYNTLTSSDEDEYEMESPNTIDTMKFFL----- 118
|||||
Db 80 YSPSLRTRLTITKTDSKNQVLTMTNVDPADTATYTCGYSVEGYGQGRPHSGQGLVT 139

QY 119 -----YVDKTHTCPPCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPE 172
|||||
Db 140 VSSEPKSCDTHTCPPCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPE 199

QY 173 VKFNWYDGVGVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPI 232
|||||
Db 200 VKFNWYDGVGVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPI 259

A:Reference number: A93157; MUID:69064124; PMID:5782707
A:Contents: annotation; Sa, disulfide bonds
C:Genetics:

A:Gene: GDB:IGHG2

A:Cross-references: GDB:119338; OMIM:147110

A:Map position: 14q32.33-14q32.33

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger disulfide bonds. Immunoglobulin C region; immunoglobulin homology
C:Superfamily: immunoglobulin C region; heterotetramer; immunoglobulin
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin homology <IM2>
F:133-202/Domain: immunoglobulin homology <IM2>
F:20-85/Domain: immunoglobulin homology <IM3>

F:14/Disulfide bonds: interchain (to light chain) #status experimental

F:17/Disulfide bonds: interchain (to heavy chain) #status experimental

F:102,103,106,109/Disulfide bonds: interchain (to heavy chain) #status experimental

F:176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 61.7%; Score 1145.5; DB 1; Length 326;
Best Local Similarity 91.0%; Pred. No. 3.3e-74;

Matches 213; Conservative 8; Mismatches 6; Indels 7; Gaps 2;

QY 120 VDKT-----HTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEV 173

Db 94 VDKTVERKCCVCCPCAPP-VAGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEV 152

QY 174 FENYVDGVEVHNAKTPREQYNSTYRVSVLTVLHQDWLNGKEYCKVSNKALPAPIE 233

Db 153 QFNWYVDGVEVHNAKTPREQNFSTFRVSVLTVLHQDWLNGKEYCKVSNKGLPAPIE 212

QY 234 KTISKAGQREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKT 293

Db 213 KTISKAGQREPQVYTLPPSRDEETKKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKT 272

QY 294 TTPVLDSGSGFFLYSLKTVDSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 347

Db 273 TTPMLDSGSGFFLYSLKTVDSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 326

RESULT 7

A60764

Ig gamma-3 chain C region, form LAT - human

C:Species: Homo sapiens (man)

C>Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 16-Jul-1999

C:Accession: A60764

R:Huck, S.; Lefranc, G.; Lefranc, M.P.

Immunogenetics 30, 250-257, 1989

A:Title: A human immunoglobulin IGHG3 allele (Gmb0, b1, c3, c5, u) with an IGHG4 convert

A:Reference number: A60764; MUID:90007613; PMID:2571587

A:Status: preliminary

A:Molecule type: DNA

A:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: immunoglobulin

F:20-85/Domain: immunoglobulin homology <IM>

Query Match 61.6%; Score 1144; DB 2; Length 377;

Best Local Similarity 92.5%; Pred. No. 5e-74;

Matches 210; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 121 DKRTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVFNMYVD 180

Db 151 DTPPPCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQKMYVD 210

QY 181 GVEVHNAKTPREQYNSTYRVSVLTVLHQDWLNGKEYCKVSNKALPAPIETISKAK 240

Db 211 GVEVHNAKTPREQYNSTYRVSVLTVLHQDWLNGKEYCKVSNKALPAPIETISKKT 270

QY 241 GQREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 300

Db 271 GQREPQVYTLPPSRDEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYTTPPVLD 330

QY 301 DGSFFLYSKLVDSKRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 347

Db 331 DGSFFLYSRLTVDSKRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 377

RESULT 8

G4HU

Ig gamma-4 chain C region - human

C:Species: Homo sapiens (man)

C>Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 16-Jul-1999

C:Accession: A90933; A90249; A02150

R:Ellison, J.; Buxbaum, J.; Hood, L.

DNA 1, 11-181, 1981

A:Title: Nucleotide sequence of a human immunoglobulin C-gamma4 gene.

A:Reference number: A90933; MUID:83157104; PMID:6299662

A:Accession: A90933

A:Molecule type: DNA

A:Residues: 1-327 <ELL>

A:Note: the sequence was determined from the germline gene

R:Pink, J.R.L.; Buttery, S.H.; De Vries, G.M.; Milstein, C.

Biochem. J. 117, 33-47, 1970

A:Title: Human immunoglobulin subclases. Partial amino acid sequence of the constan

A:Reference number: A90249; MUID:70207560; PMID:4192699

A:Accession: A90249

A:Molecule type: protein

A:Residues: 1-30; 81-326 <PIN>

C:Genetics:

A:Gene: GDB:IGHG4

A:Cross-references: GDB:119340; OMIM:147130

A:Map position: 14q32.33-14q32.33

A:Introns: 99/1; 111/1; 221/1

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (

chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin

F:20-85/Domain: immunoglobulin homology <IM1>

F:99-110/Region: hinge

F:134-203/Domain: immunoglobulin homology <IM2>

F:240-307/Domain: immunoglobulin homology <IM3>

F:14/Disulfide bonds: interchain (to light chain) #status experimental

F:27-83,141-201,247-305/Disulfide bonds: #status predicted

F:106,109/Disulfide bonds: interchain (to heavy chain) #status experimental

F:177/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 61.2%; Score 1135; DB 1; Length 327;

Best Local Similarity 93.7%; Pred. No. 1.8e-73;

Matches 208; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 126 CPGCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVH 185

Db 106 CPSCPAPEFLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFNMYVDGVEVH 165

QY 186 NAKTKPREQYNSTYRVSVLTVLHQDWLNGKEYCKVSNKALPAPIETISKAGOPRE 245

Db 166 NAKTKPREQFNSTYRVSVLTVLHQDWLNGKEYCKVSNKGLPSSIEKTISKAGOPRE 225

QY 246 POVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFF 305

Db 226 POVYTLPPSDEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFF 285

QY 306 LYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 347

Db 286 LYSRLTVDSKRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 327

RESULT 9

G3HUW1

Ig gamma-3 heavy chain disease proteins - human

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1979 #sequence_revision 23-Oct-1981 #text_change 16-Jul-1999

C:Accession: A90442; A92219; A90198; A93915; A02149

R:Frangione, B.; Rosenwasser, E.; Prelli, F.; Franklin, E.C.

Biochemistry 19, 4304-4308, 1980

A:Title: Primary structure of human gamma3 immunoglobulin deletion mutant: gamma3 heavy
 A:Reference number: A90442; MUID:81021548; PMID:6774747
 A:Contents: heavy chain disease protein Wis
 A:Accession: A90442
 A:Molecule type: protein
 A:Residues: 1-289 <FRA>
 A:Note: the molecule is a dimer linked by 12 disulfide bonds: it has an extra interchain
 A:Note: this protein lacks most of the V region and all of the CH1 region. Residue 12 co
 A:Note: the sequence of residues 42-76 was taken from the reference that follows
 R:Michaelson, T.E.; Frangione, B.; Franklin, E.C.
 J. Biol. Chem. 252, 883-889, 1977
 A:Title: Primary structure of the 'hinge' region of human IgG3. Probable quadruplication
 A:Reference number: A92219; MUID:77118561; PMID:402363
 A:Contents: normal gamma-3 chains, sequence corresponding to residues 12-97 of protein W
 A:Accession: A92219
 A:Molecule type: protein
 A:Residues: 12-97 <MIC>
 A:Note: the hinge region in gamma-3 chains is about four times as long as in other gamma
 A:Note: cysteines at positions 24, 27, 33, 39, 42, 48, 54, 57, 63, 69, and 72 form inter
 R:Wolfenstein-Todel, C.; Frangione, B.; Prelli, F.; Franklin, E.C.
 Biochem. Biophys. Res. Commun. 71, 907-914, 1976
 A:Title: The amino acid sequence of "heavy chain disease" protein ZUC. Structure of the
 A:Reference number: A90198; MUID:77021516; PMID:823945
 A:Contents: heavy chain disease protein zuc, partial sequence corresponding to residues
 A:Accession: A90198
 A:Molecule type: protein
 A:Residues: 59-125, 'BB', 128-226, 228-289 <WOL>
 A:Note: this protein lacks most of the V region, all of the CH1 region, and part of the
 R:Alexander, A.; Steinmetz, M.; Barritault, D.; Frangione, B.; Franklin, E.C.; Hood, L.;
 Proc. Natl. Acad. Sci. U.S.A. 79, 3260-3264, 1982
 A:Title: gamma heavy chain disease in man: cDNA sequence supports partial gene deletion
 A:Reference number: A93915; MUID:82247835; PMID:6808505
 A:Contents: heavy chain disease protein Omm
 A:Accession: A93915
 A:Molecule type: mRNA
 A:Residues: 12-70, 72-114, 116-125, 'E', 127-133, 'L', 135-136, 'E', 138, 'Y', 140-154, 'D', 156-157
 A:Note: a carboxyl-terminal Lys is removed posttranslationally
 A:Note: this sequence may represent an allelic form or another gamma chain subclass
 C:Comment: The heavy chain disease protein Wis is shown.
 C:Genetics:
 A:Gene: GDB:IGHG3
 A:Cross-references: GDB:119339; OMIM:147120
 A:Map position: 14q32.33-14q32.33
 C:Superfamily: immunoglobulin C region; immunoglobulin; pyroglutamic acid
 C:Keywords: duplication; glycoprotein; immunoglobulin; pyroglutamic acid
 F:203-270/Domain: immunoglobulin homology <IMW>
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
 F:6,140/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 60.4%; Score 1121; DB 1; Length 289;
 Best Local Similarity 90.3%; Pred. No. 1.6e-72;
 Matches 204; Conservative 13; Mismatches 9; Indels 0; Gaps 0;

Qy 121 DKTHCTPCPAPELLGGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 180
 Db 64 DTPPCPCPAPELLGGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 123
 Qy 181 GVEVHNATKPREQYNSTRVSVLTVLHQDLNGKEYCKVSNKALPAPIEKTISKAK 240
 Db 124 GVQVHNATKPREQYNSTRVSVLTVLHQDLNGKEYCKVSNKALPAPIEKTISKAK 183
 Qy 241 GQPREQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEVNESGQPPENNYKTTTPVLDS 300
 Db 184 GQPREQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEVNESGQPPENNYKTTTPVLDS 243
 Qy 301 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPG 346
 Db 244 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPG 289

RESULT 10
 GHRB

Ig gamma chain C region ~ rabbit
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 24-Apr-1984 #sequence,revision 15-Nov-1984 #text_change 16-Jul-1999
 C:Accession: A91749; A90290; A93928; A90245; A94416; A02161
 R:Bernstein, K.E.; Alexander, C.B.; Mage, R.G.
 Immunogenetics 18, 387-397, 1983
 A:Title: Nucleotide sequence of a rabbit IgG heavy chain from the recombinant F-I hap
 A:Reference number: A91749; MUID:84030930; PMID:6313520
 A:Accession: A91749
 A:Molecule type: mRNA
 A:Residues: 1-323 <BER>
 A:Note: this sequence has the d12 allotypic marker, 104-Thr, and the e14 marker, 185-
 R:Pratt, D.M.; Mole, L.E.
 Biochem. J. 151, 337-349, 1975
 A:Title: Sequence studies on the constant region of the Fd sections of rabbit immunog
 A:Reference number: A90290; MUID:76135469; PMID:1243651
 A:Accession: A90290
 A:Molecule type: protein
 A:Residues: 1-47, 'E', 49-71, 'PV', 72-128 <PRA>
 R:Martens, C.L.; Moore, K.W.; Steinmetz, M.; Hood, L.; Knight, K.L.
 Proc. Natl. Acad. Sci. U.S.A. 79, 6018-6022, 1982
 A:Title: Heavy chain genes of rabbit IgG: isolation of a cDNA encoding gamma heavy ch
 A:Reference number: A93928; MUID:83299917; PMID:6193512
 A:Accession: A93928
 A:Molecule type: mRNA
 A:Residues: 88-103, 'M', 105-143, 'E', 145-184, 'A', 186, 'E', 188-266 <MAR>
 A:Cross-references: GB:ML6426; NID:gl65111; PIDN:AA31289.1; PID:gl65112
 A:Note: this sequence has the d11 allotypic marker, 104-Met, and the e15 allotypic ma
 R:Fruchter, R.G.; Jackson, S.A.; Mole, L.E.; Porter, R.R.
 Biochem. J. 116, 249-259, 1970
 A:Title: Sequence studies of the Fd section of the heavy chain of rabbit immunoglobul
 A:Reference number: A90245; MUID:70110015; PMID:5461106
 A:Accession: A90245
 A:Molecule type: protein
 A:Residues: 132-143, 'E', 145-161 <FRU>
 R:Hill, R.L.; Lebovitz, H.E.; Fellows Jr., R.E.; Delaney, R.
 In Gamma Globulins, Nobel Symp. 3, Killander, J., ed., pp.109-127, Almquist and Wikse
 A:Reference number: A94416
 A:Accession: A94416
 A:Molecule type: protein
 A:Residues: 129-131;155-172, 'D', 174-184, 'A', 186, 'E', 188-200, 'D', 202-217, 'E', 219-232, '
 A:Note: this has the e15 allotypic marker, 185-Ala
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (
 hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into
 C:Superfamily: immunoglobulin C region; immunoglobulin homology
 C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
 F:20-82/Domain: immunoglobulin homology <IM1>
 F:130-199/Domain: immunoglobulin homology <IM2>
 F:236-303/Domain: immunoglobulin homology <IM3>
 F:173/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 49.9%; Score 926.5; DB 1; Length 323;
 Best Local Similarity 63.3%; Pred. No. 1.2e-58;
 Matches 178; Conservative 38; Mismatches 52; Indels 13; Gaps 4;

Qy 72 RASFSEKRVYLDTVSGSLTIYNTSSDEDEEMESPNTDTMKFFLYVDKT---HTC-- 126
 Db 51 RTFSPVRSGLYSLS----SVSVTSQSPVCNVAHATNWK-----VDKTVAPSTCSK 102
 Qy 127 PCPCAPPELLGGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHN 186
 Db 103 PTCPPPELLGGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFTWYINNEQVRT 162
 Qy 187 AKTKPREQYNSTRVSVLTVLHQDLNGKEYCKVSNKALPAPIEKTISKARGQPREP 246
 Db 163 APRPLEQGFNSTIRVSTLPTIHQDLNGKEYCKVSNKALPAPIEKTISKARGQPLEP 222
 Qy 247 QVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEVNESGQPPENNYKTTTPVLDSGSPFL 306
 Db 223 KVTMGPPREELSSRSVSLTCWINGFYPSDISVEKNGKAEDNFKTTPAVLSDGSPFL 282
 Qy 307 YSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPG 347

C>Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C/Accession: I47158

R/Kacskovics, I.; Sun, J.; Butler, J.E.

J. Immunol. 153, 3565-3573, 1994

A>Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a

A/Reference number: I47158; MUID:95015845; PMID:7930579

A/Accession: I47158

A/Status: preliminary; translated from GB/EMBL/DDBJ

A/Molecule type: mRNA

A/Residues: 1-328 <KAC>

A/Cross-references: EMBL:U03778; NID:g433121; PID:AAA52216.1; PID:g433122

C/Genetics:

A/Gene: IgG1

C/Superfamily: immunoglobulin C region: immunoglobulin homology

F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 48.2%; Score 894.5; DB 2; Length 328;

Best Local Similarity 58.8%; Pred. No. 2.4e-56;

Matches 177; Conservative 37; Mismatches 60; Indels 27; Gaps 4;

QY 73 AFSSFKNRVLDIVTSGSLT-----IYNLT-----SSDEDEYEMESP 108

DB 29 ASSYFPEVTVTNWNGALTSVHTPPSVLPQSGLYSLSMVTVPASSLSKSYTCNVNHP 88

QY 109 NITDMKFFLYVDKTHTCPCAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDYSH 168

DB 89 ATTTKVKRGIHQPTCPICGCE-VAGPSVFIFFPKPKDTLMISQTPVTCVVVDVSK 147

QY 169 EDEPKFNWYVDGVVHNAKTPREEQYNSTYRVSVLTVLQDMLNGKEYKCKVSNKAL 228

DB 148 EHAEOFSWYVDGVVHNAETRPKEEQYNSTYRVSVLPIQDMLNGKEYKCKVSNVNDL 207

QY 229 PAPIEKTISKAKGPPEQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAYEWESNGO-- 286

DB 208 PAPITRIKAIQSGREPQVYTLPPAEELSKSVTLTCLVIGFPPDIHWNKSNQPE 267

QY 287 PENNYKTPPVLDSDGSFFLYSKLVYDKSRWQGNVFCSCVMHEALHNNHYTKSLSPG 346

DB 268 PENTYRTPTPPQDDVDTFFLYSKLVADKRDHGDGKFCVAMHEALHNNHYTKSLSTQG 327

QY 347 K 347

DB 328 K 328

RESULT 15

G2BP

Ig gamma-2 chain C region - guinea pig

C/Species: Cavia porcellus (guinea pig)

C/Date: 07-May-1981 #sequence_revision 07-May-1981 #text_change 16-Jul-1999

C/Accession: A9453; A90352; A90359; A90384; A90385; A02151

R/Trischmann, T.M.

submitted to the Atlas, April 1975

A/Reference number: A94553

A/Molecule type: protein

A/Residues: 1-3 <TR>

Biochemistry 10, 18-25, 1971

A/Title: Structure of heavy chain from strain 13 guinea pig immunoglobulin-G(2). III. Am

A/Reference number: A90352; MUID:71058471; PMID:5538606

A/Accession: A90352

A/Molecule type: protein

A/Residues: 4-68 <BTR>

R/Turner, K.J.; Cebra, J.J.

Biochemistry 10, 9-17, 1971

A/Title: Structure of heavy chain from strain 13 guinea pig immunoglobulin-G(2). II. Am

A/Reference number: A90359; MUID:71058486; PMID:5538616

A/Accession: A90359

A/Molecule type: protein

A/Residues: 69-133;312-329 <TUR>

R/Tracey, D.E.; Cebra, J.J.

Biochemistry 13, 4796-4803, 1974

A>Title: Primary structure of the C-H2 homology region from guinea pig IgG2 antibody

A/Reference number: A90384; MUID:75036072; PMID:4429665

A/Accession: A90384

A/Molecule type: protein

A/Residues: 134-226 <TRA>

R/Trischmann, T.M.; Cebra, J.J.

Biochemistry 13, 4804-4811, 1974

A>Title: Primary structure of the C-H3 homology region from guinea pig IgG2 antibody

A/Reference number: A90385; MUID:75036073; PMID:4609467

A/Accession: A90385

A/Molecule type: protein

A/Residues: 227-311 <TR2>

R/Oliveira, B.; Lamm, M.E.

Biochemistry 10, 26-31, 1971

A>Title: Interchain disulfide bridges of guinea pig gamma-2- immunoglobulin.

A/Reference number: A90354; MUID:71058474; PMID:4922544

A/Contents: annotation; disulfide bonds

A/Note: Cys-16 is involved in a heavy-light chain bond

C/Comment: This chain was isolated from pooled serum of strain 13 inbred guinea pigs

C/Complex: An immunoglobulin heterotrimer subunit consists of two identical light (

chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into

C/Superfamily: immunoglobulin C region: immunoglobulin homology

C/Keywords: duplication; glycoprotein; heterotrimer; immunoglobulin

F:135-204/Domain: immunoglobulin homology <IM2>

F:241-310/Domain: immunoglobulin homology <IM3>

F:28-79/Disulfide bonds: #status experimental

F:142-202/Disulfide bonds: #status experimental

F:178/Binding site: carbohydrate (Asn) (covalent) #status experimental

F:248-308/Disulfide bonds: #status experimental

Query Match 48.1%; Score 892; DB 1; Length 329;

Best Local Similarity 60.6%; Pred. No. 3.6e-56;

Matches 177; Conservative 30; Mismatches 47; Indels 38; Gaps 6;

QY 87 SGSLT-----IYNLT-----SSDEDEYEMESPNTIDTMKFFLYVDKT---- 123

DB 44 SGALTSVGHVHTPPAVLQSGLYSLTSMVTVPSSOKATCNVHP--ASSTK----VDKTVETI 97

QY 124 -----HTCCPCAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVFN 176

DB 98 RTPZBPCTCPKCPPENLGGPSVFIFFPKPKDTLMISLTPTVTCVVVDVSDPEVQFT 157

QY 177 WYVDGVVHNAKTPREEQYNSTYRVSVLTVLQDMLNGKEYKCKVSNKALPADIETI 236

DB 158 WFDNKPVGNAETKPRVEQYNTTFRVESVLPDQDMLNGKEYKCKVSNKALPADIETI 217

QY 237 SKAGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAYEWESNGQP--ENNYKTT 294

DB 218 SKTKGAPRMPDYTLPPSRDELTKNOVSLTCLVKGFPSPDIAYEWESNGQP--ENNYKTT 277

QY 295 PPVLSDGSFFLYSKLVYDKSRWQGNVFCSCVMHEALHNNHYTKSLSPG 346

DB 278 PPVEDAGSYFLYSLKLVYDKSAMQDQYTYTCSVMHEALHNNHYTKSLSPG 329

Search completed: January 28, 2003, 08:40:51

Job time : 19.0184 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 08:38:45 : Search time 11.3065 Seconds
(without alignments)
650.573 Million cell updates/sec

Title: US-09-730-465-2

Perfect score: 1326

Sequence: 1 MVAGSDAGRALGVLSVCLL.....VLYMNGILKDRKPDRNTSN 250

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2.6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2.6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2.6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2.6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2.6/ptodata/1/iaa/PCTUS_COMB.pep.*

6: /cgn2.6/ptodata/1/iaa/PCTUS_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1326	100.0	250	1	US-07-940-861-10
2	1326	100.0	250	1	US-08-459-512-10
3	1326	100.0	250	2	US-08-459-657-10
4	1326	100.0	250	2	US-08-460-132-10
5	1326	100.0	250	4	US-08-466-465-2
6	1326	100.0	250	5	PCT-US92-02050-10
7	1326	100.0	250	6	5223394-1
8	1307.5	98.6	251	6	5185441-38
9	1250	94.3	240	1	US-07-940-861-12
10	1250	94.3	240	1	US-08-459-512-12
11	1250	94.3	240	2	US-08-459-657-12
12	1250	94.3	240	2	US-08-460-132-12
13	1250	94.3	240	4	US-08-466-465-4
14	1250	94.3	240	5	PCT-US92-02050-12
15	1250	94.3	240	6	5185441-36
16	1250	94.3	240	6	5223394-4
17	1250	94.3	240	6	5223394-6
18	1179	88.9	222	1	US-08-328-152A-8
19	647	48.8	134	1	US-08-328-152A-13
20	630	47.5	347	1	US-07-940-861-43
21	630	47.5	347	1	US-08-459-512-43
22	630	47.5	347	2	US-08-459-657-43
23	630	47.5	347	2	US-08-460-132-43
24	630	47.5	347	4	US-08-466-465-8
25	630	47.5	347	5	PCT-US92-02050-43
26	471.5	35.6	225	1	US-08-328-152A-31
27	447.5	33.7	199	1	US-08-328-152A-36

28	402	30.3	77	1	US-07-940-861-5	Sequence 5, Appli
29	402	30.3	77	1	US-08-459-512-5	Sequence 5, Appli
30	402	30.3	77	2	US-08-459-657-5	Sequence 5, Appli
31	402	30.3	77	2	US-08-460-132-5	Sequence 5, Appli
32	402	30.3	77	5	PCT-US92-02050-5	Sequence 5, Appli
33	281	21.2	318	6	5223394-11	Patent No. 5223394
34	260	19.6	50	1	US-07-940-861-2	Sequence 2, Appli
35	260	19.6	50	1	US-08-459-512-2	Sequence 2, Appli
36	260	19.6	50	2	US-08-459-657-2	Sequence 2, Appli
37	260	19.6	50	2	US-08-460-132-2	Sequence 2, Appli
38	260	19.6	50	5	PCT-US92-02050-2	Sequence 2, Appli
39	256.5	19.3	131	1	US-08-328-152A-1	Sequence 1, Appli
40	200	15.1	38	6	5185441-1	Patent No. 5185441
41	165	12.4	32	1	US-07-940-861-33	Sequence 33, Appl
42	165	12.4	32	1	US-08-459-512-33	Sequence 33, Appl
43	165	12.4	32	2	US-08-459-657-33	Sequence 33, Appl
44	165	12.4	32	2	US-08-460-132-33	Sequence 33, Appl
45	165	12.4	32	5	PCT-US92-02050-33	Sequence 33, Appl

ALIGNMENTS

RESULT 1
US-07-940-861-10
: Sequence 10, Application US/07940861
: Patent No. 5547853
: GENERAL INFORMATION:
: APPLICANT: BIOGEN, INC.
: APPLICANT: WALLNER, Barbara P.
: APPLICANT: MILLER, Glenn T.
: APPLICANT: ROSA, Margaret D.
: TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
: TITLE OF INVENTION: FUNCTION ASSOCIATED ANTIGEN 3
: NUMBER OF SEQUENCES: 43
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Neave
: STREET: 875 Third Avenue
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10022-6250
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/940,861
: FILING DATE: 21-OCT-1992
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/US92/02050
: FILING DATE: 12-MAR-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/667,971
: FILING DATE: 12-MAR-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/770,967
: FILING DATE: 07-OCT-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: HALEY, James F., Jr.
: REGISTRATION NUMBER: 27,794
: REFERENCE/DOCKET NUMBER: B151C1P2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212)715-0600
: TELEFAX: (212)715-0673
: TELEX: 14-8367
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 250 amino acids
: TYPE: amino acid
: TOPOLOGY: linear

```
; MOLECULE TYPE: protein
US-07-940-861-10

Query Match      100.0%; Score 1326; DB 1; Length 250;
Best Local Similarity 100.0%; Pred. NO. 1.7e-126;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVAGSDAGRALGVLSVCLLCGFCFISCFQQIYGVVYGNVTFHVPNSNPLKEVLKKQK 60
   |||||
Db 1 MVAGSDAGRALGVLSVCLLCGFCFISCFQQIYGVVYGNVTFHVPNSNPLKEVLKKQK 60

QY 61 DKVAELENSEFRASFKNRYVLDTVSGSLTIYNTSSDEDEYEMESPNITDTMKFFLYV 120
   |||||
Db 61 DKVAELENSEFRASFKNRYVLDTVSGSLTIYNTSSDEDEYEMESPNITDTMKFFLYV 120

QY 121 LESLPSPTLTCAITNGSIEVQCMIPSHYNSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
   |||||
Db 121 LESLPSPTLTCAITNGSIEVQCMIPSHYNSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180

QY 181 LPQKIOCTLSNPLNTSSIIITTCIPSSGHSRHRVALIPIPLAVITTCIVLYMNGILKC 240
   |||||
Db 181 LPQKIOCTLSNPLNTSSIIITTCIPSSGHSRHRVALIPIPLAVITTCIVLYMNGILKC 240

QY 241 DRKPDRTNSN 250
   |||||
Db 241 DRKPDRTNSN 250

RESULT 2
US-08-459-512-10
; Sequence 10, Application US/08459512
; Patent No. 5728677
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: WALLNER, Barbara P.
; APPLICANT: MILLER, Glenn T.
; APPLICANT: ROSA, Margaret D.
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 875 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022-6250
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,512
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02050
; FILING DATE: 12-MAR-1992
; APPLICATION NUMBER: US 07/667,971
; FILING DATE: 12-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,967
; FILING DATE: 07-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY, James F., Jr.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B151CIP2
; TELEPHONE: (212)715-0600
; TELEFAX: (212)715-0673
; TELEEX: 14-8367
; INFORMATION FOR SEQ ID NO: 10:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-459-512-10

Query Match      100.0%; Score 1326; DB 1; Length 250;
Best Local Similarity 100.0%; Pred. NO. 1.7e-126;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVAGSDAGRALGVLSVCLLCGFCFISCFQQIYGVVYGNVTFHVPNSNPLKEVLKKQK 60
   |||||
Db 1 MVAGSDAGRALGVLSVCLLCGFCFISCFQQIYGVVYGNVTFHVPNSNPLKEVLKKQK 60

QY 61 DKVAELENSEFRASFKNRYVLDTVSGSLTIYNTSSDEDEYEMESPNITDTMKFFLYV 120
   |||||
Db 61 DKVAELENSEFRASFKNRYVLDTVSGSLTIYNTSSDEDEYEMESPNITDTMKFFLYV 120

QY 121 LESLPSPTLTCAITNGSIEVQCMIPSHYNSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
   |||||
Db 121 LESLPSPTLTCAITNGSIEVQCMIPSHYNSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180

QY 181 LPQKIOCTLSNPLNTSSIIITTCIPSSGHSRHRVALIPIPLAVITTCIVLYMNGILKC 240
   |||||
Db 181 LPQKIOCTLSNPLNTSSIIITTCIPSSGHSRHRVALIPIPLAVITTCIVLYMNGILKC 240

QY 241 DRKPDRTNSN 250
   |||||
Db 241 DRKPDRTNSN 250

RESULT 3
US-08-459-657-10
; Sequence 10, Application US/08459657
; Patent No. 5914111
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: WALLNER, Barbara P.
; APPLICANT: MILLER, Glenn T.
; APPLICANT: ROSA, Margaret D.
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 875 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022-6250
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,657
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02050
; FILING DATE: 12-MAR-1992
; APPLICATION NUMBER: US 07/667,971
; FILING DATE: 12-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,967
; FILING DATE: 07-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY, James F., Jr.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B151CIP2
; TELEPHONE: (212)715-0600
; TELEFAX: (212)715-0673
; TELEEX: 14-8367
; INFORMATION FOR SEQ ID NO: 10:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-459-512-10
```


QY 121 LESLPSPTLTCAITNGSIEVOCMIPEHYNHSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
|
Db 121 LESLPSPTLTCAITNGSIEVOCMIPEHYNHSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
|
QY 181 LPQKIOCTLSNPLFNTSSIIILTTCTIPSSGHSRHRALPIPLAVITTCIVLYMNGILKC 240
|
Db 181 LPQKIOCTLSNPLFNTSSIIILTTCTIPSSGHSRHRALPIPLAVITTCIVLYMNGILKC 240
|
QY 241 DRKPDRTNSN 250
|
Db 241 DRKPDRTNSN 250
|

RESULT 8

5185441-38
; Patent No. 5185441
; APPLICANT: WALLNER, BARBARA P.; HESSELS, CATHERINE
; TITLE OF INVENTION: DNA SEQUENCES, RECOMBINANT DNA
; MOLECULES AND PROCESSES FOR PRODUCING PI-LINKED LYMPHOCYTE
; FUNCTION ASSOCIATED ANTIGEN-3
; NUMBER OF SEQUENCES: 41
; CURRENT APPLICATION NUMBER: US/07/237,309
; FILING DATE: 26-AUG-1988
; SEQ ID NO: 38;
; LENGTH: 251
5185441-38

Query Match 98.6%; Score 1307.5; DB 6; Length 251;
Best Local Similarity 99.2%; Pred. No. 1.3e-124;
Matches 249; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 MVAGSDAGRALGVLSVWCLLHCFGFISCFSQIYGVVGNVTFHVPNSVPLKEVLWKQK 60
|
Db 1 MVAGSDAGRALGVLSVWCLLHCFGFISCFSQIYGVVGNVTFHVPNSVPLKEVLWKQK 60
|
QY 61 DKVAELENSEFRASFKNRVYLDTVSGSLTIYNLTSSDEDEYEMESPNITDTMKFFLYV 120
|
Db 61 DKVAELENSEFRASFKNRVYLDTVSGSLTIYNLTSSDEDEYEMESPNITDTMKFFLYV 120
|
QY 121 LESLPSPTLTCAITNGSIEVOCMIPEHYNHSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
|
Db 121 LESLPSPTLTCAITNGSIEVOCMIPEHYNHSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
|
QY 181 LPQKIOCTLSNPLFNTSSIIILTTCTIPSSGHSRHRALPIPLAVITTCIVLYMNG-ILK 239
|
Db 181 LPQKIOCTLSNPLFNTSSIIILTTCTIPSSGHSRHRALPIPLAVITTCIVLYMNGILK 240
|
QY 240 DRKPDRTNSN 250
|
Db 241 DRKPDRTNSN 251
|

RESULT 9

US-07-940-861-12
; Sequence 12, Application US/07940861
; Patent No. 5547853
; GENERAL INFORMATION:

; APPLICANT: BIOGEN, INC.
; APPLICANT: WALLNER, BARBARA P.
; APPLICANT: MILLER, GLENN T.
; APPLICANT: ROSA, MARGARET D.
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 875 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022-6250
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/940,861
; FILING DATE: 21-OCT-1992
; CLASSIFICATION: 514

PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US92/02050

; FILING DATE: 12-MAR-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/667,971

; FILING DATE: 12-MAR-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/770,967

; FILING DATE: 07-OCT-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: HALEY, JAMES F., JR.

; REGISTRATION NUMBER: 27,794

; REFERENCE/DOCKET NUMBER: B51CIP2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 715-0600

; TELEFAX: (212) 715-0673

; TELEX: 14-8367

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 240 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-07-940-861-12

Query Match 94.3%; Score 1250; DB 1; Length 240;

Best Local Similarity 99.6%; Pred. No. 8.3e-119;

Matches 236; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVAGSDAGRALGVLSVWCLLHCFGFISCFSQIYGVVGNVTFHVPNSVPLKEVLWKQK 60

|

Db 1 MVAGSDAGRALGVLSVWCLLHCFGFISCFSQIYGVVGNVTFHVPNSVPLKEVLWKQK 60

|

QY 61 DKVAELENSEFRASFKNRVYLDTVSGSLTIYNLTSSDEDEYEMESPNITDTMKFFLYV 120

|

Db 61 DKVAELENSEFRASFKNRVYLDTVSGSLTIYNLTSSDEDEYEMESPNITDTMKFFLYV 120

|

QY 121 LESLPSPTLTCAITNGSIEVOCMIPEHYNHSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180

|

Db 121 LESLPSPTLTCAITNGSIEVOCMIPEHYNHSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180

|

QY 181 LPQKIOCTLSNPLFNTSSIIILTTCTIPSSGHSRHRALPIPLAVITTCIVLYMNGI 237

|

Db 181 LPQKIOCTLSNPLFNTSSIIILTTCTIPSSGHSRHRALPIPLAVITTCIVLYMNGM 237

|

RESULT 10

US-08-459-512-12

; Sequence 12, Application US/08459512

; Patent No. 5728677

; GENERAL INFORMATION:

; APPLICANT: BIOGEN, INC.

; APPLICANT: WALLNER, BARBARA P.

; APPLICANT: MILLER, GLENN T.

; APPLICANT: ROSA, MARGARET D.

; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE

; NUMBER OF SEQUENCES: 43

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Neave

; STREET: 875 Third Avenue

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10022-6250

; COMPUTER READABLE FORM:

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,512
FILING DATE: 02-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/02050
FILING DATE: 12-MAR-1992
APPLICATION NUMBER: US 07/667,971
FILING DATE: 12-MAR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/770,967
FILING DATE: 07-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: HALEY, James F., JR.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B151CIP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)715-0600
TELEFAX: (212)715-0673
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 240 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-459-512-12

Query Match 94.3%; Score 1250; DB 1; Length 240;
Best Local Similarity 99.6%; Pred. No. 8.3e-119;
Matches 236; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWAGSDAGRALGVLSVWCLLCFCGFTSCFSQQIYGVVYGNVTFHVPNSVPLKEVLKKOK 60
DB 1 MWAGSDAGRALGVLSVWCLLCFCGFTSCFSQQIYGVVYGNVTFHVPNSVPLKEVLKKOK 60
QY 61 DKVAELENSEFRASFKNRYLDTVSGSLTIYNTLTSSDEDEYEMESPNTIDTMKFFLV 120
DB 61 DKVAELENSEFRASFKNRYLDTVSGSLTIYNTLTSSDEDEYEMESPNTIDTMKFFLV 120
QY 121 LESLPSTLTCAITNGSTEVOCMIPHYNSHRGLIMYSWDCPMEOCKRNSTSIYFKMEND 180
DB 121 LESLPSTLTCAITNGSTEVOCMIPHYNSHRGLIMYSWDCPMEOCKRNSTSIYFKMEND 180
QY 181 LPQKIOCTLSNPLFNWTTSSIIITTCIPSSGHSRHRVALIPIPLAVITTCIVLYMNGI 237
DB 181 LPQKIOCTLSNPLFNWTTSSIIITTCIPSSGHSRHRVALIPIPLAVITTCIVLYMNGM 237

RESULT 11
US-08-459-657-12
Sequence 12, Application US/08459657
Patent No. 5914111
GENERAL INFORMATION:
APPLICANT: BIOGEN, INC.
APPLICANT: WALLNER, Barbara P.
APPLICANT: MILLER, Glenn T.
APPLICANT: ROSA, Margaret D.
TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Neave
STREET: 875 Third Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10022-6250

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,657
FILING DATE: 02-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/02050
FILING DATE: 12-MAR-1992
APPLICATION NUMBER: US 07/667,971
FILING DATE: 12-MAR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/770,967
FILING DATE: 07-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: HALEY, James F., JR.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B151CIP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)715-0600
TELEFAX: (212)715-0673
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 240 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-459-657-12

Query Match 94.3%; Score 1250; DB 2; Length 240;
Best Local Similarity 99.6%; Pred. No. 8.3e-119;
Matches 236; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWAGSDAGRALGVLSVWCLLCFCGFTSCFSQQIYGVVYGNVTFHVPNSVPLKEVLKKOK 60
DB 1 MWAGSDAGRALGVLSVWCLLCFCGFTSCFSQQIYGVVYGNVTFHVPNSVPLKEVLKKOK 60
QY 61 DKVAELENSEFRASFKNRYLDTVSGSLTIYNTLTSSDEDEYEMESPNTIDTMKFFLV 120
DB 61 DKVAELENSEFRASFKNRYLDTVSGSLTIYNTLTSSDEDEYEMESPNTIDTMKFFLV 120
QY 121 LESLPSTLTCAITNGSTEVOCMIPHYNSHRGLIMYSWDCPMEOCKRNSTSIYFKMEND 180
DB 121 LESLPSTLTCAITNGSTEVOCMIPHYNSHRGLIMYSWDCPMEOCKRNSTSIYFKMEND 180
QY 181 LPQKIOCTLSNPLFNWTTSSIIITTCIPSSGHSRHRVALIPIPLAVITTCIVLYMNGI 237
DB 181 LPQKIOCTLSNPLFNWTTSSIIITTCIPSSGHSRHRVALIPIPLAVITTCIVLYMNGM 237

RESULT 12
US-08-460-132-12
Sequence 12, Application US/08460132
Patent No. 5928643
GENERAL INFORMATION:
APPLICANT: BIOGEN, INC.
APPLICANT: WALLNER, Barbara P.
APPLICANT: MILLER, Glenn T.
APPLICANT: ROSA, Margaret D.
TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Neave
STREET: 875 Third Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10022-6250

```
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/460,132
;; FILING DATE: 02-JUN-1995
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/940,861
;; FILING DATE: 21-OCT-1992
;; APPLICATION NUMBER: PCT/US92/02050
;; FILING DATE: 12-MAR-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/667,971
;; FILING DATE: 12-MAR-1991
;; APPLICATION DATA:
;; APPLICATION NUMBER: US 07/770,967
;; FILING DATE: 07-OCT-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: HALEY, James F., Jr.
;; REGISTRATION NUMBER: 27,794
;; REFERENCE/DOCKET NUMBER: B151CIP2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212)715-0600
;; TELEFAX: (212)715-0673
;; TELEX: 14-8367
;; INFORMATION FOR SEQ ID NO: 12:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 240 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-460-132-12

Query Match 94.3%; Score 1250; DB 2; Length 240;
Best Local Similarity 93.6%; Pred. No. 8.3e-119;
Matches 236; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVAGSDAGRALGVLSVWCLLHCFGFSQSIYGVYGNVTFHVPNSVPLKEVLWKKQK 60
Db 1 MVAGSDAGRALGVLSVWCLLHCFGFSQSIYGVYGNVTFHVPNSVPLKEVLWKKQK 60
Qy 61 DKVAELENSEFRAFSSFKNRVYLDTVSGSLTIYNTSSDEDEYEMESPNITDTMKFFLYV 120
Db 61 DKVAELENSEFRAFSSFKNRVYLDTVSGSLTIYNTSSDEDEYEMESPNITDTMKFFLYV 120
Qy 121 LESLPSPTLTALTNGSIEVOCMIPEHYNHSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
Db 121 LESLPSPTLTALTNGSIEVOCMIPEHYNHSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
Qy 121 LESLPSPTLTALTNGSIEVOCMIPEHYNHSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
Db 121 LESLPSPTLTALTNGSIEVOCMIPEHYNHSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
Qy 181 LPQKIOCTLSNPLNFTSSIIITTCIPSSGHSRHRALYALPIPLAVITTCIVLYNMGI 237
Db 181 LPQKIOCTLSNPLNFTSSIIITTCIPSSGHSRHRALYALPIPLAVITTCIVLYNMGI 237
Db 181 LPQKIOCTLSNPLNFTSSIIITTCIPSSGHSRHRALYALPIPLAVITTCIVLYNMGM 237

RESULT 13
US-08-466-465-4
; Sequence 4, Application US/08466465
; Patent No. 6162432
; GENERAL INFORMATION:
; APPLICANT: Wallner, Barbara P.
; APPLICANT: Cooper, Kevin D.
; TITLE OF INVENTION: Method of Prophylaxis or Treatment of Antigen
; TITLE OF INVENTION: Presenting Cell Driven Skin Conditions Using
; TITLE OF INVENTION: Inhibitors of the CD2/LFA-3 Interaction
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
```

```
;; COUNTRY: USA
;; ZIP: 02109-1875
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/466,465
;; FILING DATE:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US92/08755
;; FILING DATE: 06-OCT-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/862,022
;; FILING DATE: 12-APR-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/770,969
;; FILING DATE: 07-OCT-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Myers, Louis (PLM)
;; REGISTRATION NUMBER: 35,965
;; REFERENCE/DOCKET NUMBER: BGP-111CP
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (617)227-7400
;; TELEFAX: (617)227-5941
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 240 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-466-465-4

Query Match 94.3%; Score 1250; DB 4; Length 240;
Best Local Similarity 99.6%; Pred. No. 8.3e-119;
Matches 236; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVAGSDAGRALGVLSVWCLLHCFGFSQSIYGVYGNVTFHVPNSVPLKEVLWKKQK 60
Db 1 MVAGSDAGRALGVLSVWCLLHCFGFSQSIYGVYGNVTFHVPNSVPLKEVLWKKQK 60
Qy 61 DKVAELENSEFRAFSSFKNRVYLDTVSGSLTIYNTSSDEDEYEMESPNITDTMKFFLYV 120
Db 61 DKVAELENSEFRAFSSFKNRVYLDTVSGSLTIYNTSSDEDEYEMESPNITDTMKFFLYV 120
Qy 121 LESLPSPTLTALTNGSIEVOCMIPEHYNHSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
Db 121 LESLPSPTLTALTNGSIEVOCMIPEHYNHSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
Qy 181 LPQKIOCTLSNPLNFTSSIIITTCIPSSGHSRHRALYALPIPLAVITTCIVLYNMGI 237
Db 181 LPQKIOCTLSNPLNFTSSIIITTCIPSSGHSRHRALYALPIPLAVITTCIVLYNMGM 237

RESULT 14
PCT-US92-02050-12
; Sequence 12, Application PC/TUS9202050
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: WALLNER, Barbara P.
; APPLICANT: MILLER, Glenn T.
; APPLICANT: ROSA, Margaret D.
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
; TITLE OF INVENTION: FUNCTION ASSOCIATED ANTIGEN 3
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 875 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022-6250
```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, V
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/02050
FILING DATE: 19920312
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/667,971
FILING DATE: 12-MAR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/770,967
FILING DATE: 07-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: HALEY, James F., Jr.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B51G1P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)715-0600
TELEFAX: (212)715-0673
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 240 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-02050-12

```

Query Match	94.3%	Score 1250;	DB 5;	Length 240;
Best local similarity	99.6%;	Prd. No. 8.3e-119;		
Matches 236;	Conservative	1;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	MVAGSDAGALGVLSVVCLLHCFGFI	SCFSQQIYGVVYGVNTFHVPSNVPLKEVLKKQK	60
Db	1	MVAGSDAGALGVLSVVCLLHCFGFI	SCFSQQIYGVVYGVNTFHVPSNVPLKEVLKKQK	60
QY	61	DKVAELNSBFRAFSFKNRVYLDVTSGSLTY	LNLTSSDEDEYEMESPNITDMKKFFLVX	120
Db	61	DKVAELNSBFRAFSFKNRVYLDVTSGSLTY	LNLTSSDEDEYEMESPNITDMKKFFLVX	120
QY	121	LESLSPTLTLCALNTSGIEVOCMPIEYHNHRLGLIMY	SDCPCMEQCKRNSTSYFKMEND	180
Db	121	LESLSPTLTLCALNTSGIEVOCMPIEYHNHRLGLIMY	SDCPCMEQCKRNSTSYFKMEND	180
QY	181	LPQKIQCTLSNPLNTTSSII	LTTCIPSSGSHRVALIPI	PLAVITTCIVLYMGI 237
Db	181	LPQKIQCTLSNPLNTTSSII	LTTCIPSSGSHRVALIPI	PLAVITTCIVLYMGM 237

```

RESULT 15
5185441-36
; Patent NO. 5185441-36
; APPLICANT: WALL
; TITLE OF INVENTION
; MOLECULES AND PROCESSES
; FUNCTION ASSOCIATED WITH MOLECULES
; NUMBER OF SEQUENCES
; CURRENT APPLICATION NO.
; APPLICATION DATE:
; FILING DATE:
; SEQ ID NO: 36
; LENGTH: 240
5185441-36

```

```

Query Match          94.3%; Score 1250; DB 6; Length 240;
Best Local Similarity 99.6%; Pred. No. 8.3e-119;
Matches 236; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1  MVASDAGRALGVLSVWCLLHCFCISCFQQIYGVVYGNVTFFHVPNSVPLKEVLAKKOK 60
      *  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

Db      1  MYAGSDAGRALGVLSVWCLLHCFGTSCFSQQIYGVVGNVTHFVPSNVPLKEVLWKQK 60
QY      61  DKVAELENSEFAFSSFNRRYYLDTVSGSLTIYNLTSSDEDEYEMESPNIITDTMKFFLYV 120
Db      61  DKVAELENSEFAFSSFNRRYYLDTVSGSLTIYNLTSSDEDEYEMESPNIITDTMKFFLYV 120
QY      121 LESLPSPTLTCAITNGSTSEVOCMITPEYHNSHRGLIMYSWDCPEQCKRNSTSIYFKMEND 180
Db      121 LESLPSPTLTCAITNGSTSEVOCMITPEYHNSHRGLIMYSWDCPEQCKRNSTSIYFKMEND 180
QY      181 LPQIKOICTLSNPLFNTSSIIILTCIPSSGSHRRYALIPILPAVITTCIVLYNNGI 237
Db      181 LPQIKOICTLSNPLFNTSSIIILTCIPSSGSHRRYALIPILPAVITTCIVLYNNGI 237

Search completed: January 28, 2003, 08:40:09
Job time : 13.3065 secs

```

Search completed: January 28, 2003, 08:40:09
Job time : 13.3065 secs

GenCore version 5.1.3
Copyright (C) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 08:40:59 ; Search time 82.077 Seconds
(without alignments)
61.462 Million cell updates/sec

Title: US-09-730-465-2
Perfect score: 1326
Sequence: 1 MVAGSDAGRALGVLSVCLL.....VLYMNGILCKDRKPDRTNSN 250

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues 122226

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :				Published_Applications_AA:*			
1:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*	1326	100.0	250	10	US-09-796-033-2	Sequence 2, Appli
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*	1326	100.0	250	10	US-09-730-465-2	Sequence 2, Appli
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*	1250	94.3	240	10	US-09-796-033-4	Sequence 4, Appli
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*	1250	94.3	240	10	US-09-730-465-4	Sequence 4, Appli
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*	630	47.5	347	9	US-10-091-236-17	Sequence 17, Appl
6:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*	630	47.5	347	10	US-09-796-033-8	Sequence 8, Appli
7:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*	630	47.5	347	10	US-09-730-465-8	Sequence 8, Appli
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*	115.5	8.7	351	10	US-09-796-033-6	Sequence 6, Appli
9:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*	115.5	8.7	351	10	US-09-730-465-6	Sequence 6, Appli
10:	/cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*	105.5	8.0	773	9	US-09-982-107-2	Sequence 2, Appli
11:	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*	105.5	8.0	773	10	US-09-818-247-6	Sequence 6, Appli
12:	/cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*	90	6.8	312	9	US-09-732-524-4	Sequence 4, Appli
13:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*	90	6.8	335	9	US-09-992-598-253	Sequence 253, App
14:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*	90	6.8	335	9	US-09-989-293A-253	Sequence 253, App

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1326	100.0	250	10	US-09-796-033-2 Sequence 2, Appli
2	1326	100.0	250	10	US-09-730-465-2 Sequence 2, Appli
3	1250	94.3	240	10	US-09-796-033-4 Sequence 4, Appli
4	1250	94.3	240	10	US-09-730-465-4 Sequence 4, Appli
5	630	47.5	347	9	US-10-091-236-17 Sequence 17, Appl
6	630	47.5	347	10	US-09-796-033-8 Sequence 8, Appli
7	630	47.5	347	10	US-09-730-465-8 Sequence 8, Appli
8	115.5	8.7	351	10	US-09-796-033-6 Sequence 6, Appli
9	115.5	8.7	351	10	US-09-730-465-6 Sequence 6, Appli
10	105.5	8.0	773	9	US-09-982-107-2 Sequence 2, Appli
11	105.5	8.0	773	10	US-09-818-247-6 Sequence 6, Appli
12	90	6.8	312	9	US-09-732-524-4 Sequence 4, Appli
13	90	6.8	335	9	US-09-992-598-253 Sequence 253, App
14	90	6.8	335	9	US-09-989-293A-253 Sequence 253, App
15	90	6.8	335	9	US-10-063-547-46 Sequence 46, Appl
16	90	6.8	335	9	US-09-989-735-253 Sequence 253, App
17	90	6.8	335	9	US-09-990-444-253 Sequence 253, App
18	90	6.8	335	9	US-09-989-730-253 Sequence 253, App
19	90	6.8	335	9	US-09-990-436-253 Sequence 253, App

20	90	6.8	335	9	US-09-991-181-253	Sequence 253, App
21	90	6.8	335	9	US-09-993-687-253	Sequence 253, App
22	90	6.8	335	9	US-09-989-734-253	Sequence 253, App
23	90	6.8	335	9	US-09-997-653-253	Sequence 253, App
24	90	6.8	335	9	US-10-174-590-192	Sequence 192, App
25	90	6.8	335	9	US-10-176-758-192	Sequence 192, App
26	90	6.8	335	9	US-10-063-616-46	Sequence 46, Appl
27	90	6.8	335	9	US-10-175-737-192	Sequence 192, App
28	90	6.8	335	10	US-09-732-524-2	Sequence 2, Appli
29	90	6.8	335	10	US-09-989-722-253	Sequence 253, App
30	90	6.8	335	10	US-09-989-723-253	Sequence 253, App
31	90	6.8	335	10	US-09-989-729-253	Sequence 253, App
32	90	6.8	335	10	US-09-989-727-253	Sequence 253, App
33	90	6.8	335	10	US-09-989-731-253	Sequence 253, App
34	90	6.8	335	10	US-09-989-732-253	Sequence 253, App
35	90	6.8	335	10	US-09-745-605-4	Sequence 4, Appli
36	90	6.8	335	10	US-09-991-073-253	Sequence 253, App
37	90	6.8	335	10	US-09-990-442-253	Sequence 253, App
38	90	6.8	335	10	US-09-991-163-253	Sequence 253, App
39	90	6.8	335	10	US-09-993-604-253	Sequence 253, App
40	90	6.8	335	10	US-09-990-456-253	Sequence 253, App
41	90	6.8	335	10	US-09-989-721-253	Sequence 253, App
42	90	6.8	335	12	US-10-006-867-46	Sequence 46, Appl
43	90	6.8	335	12	US-10-052-586-192	Sequence 192, App
44	88.5	6.7	306	10	US-09-837-867A-17	Sequence 17, Appl
45	88	6.6	313	10	US-09-756-983-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-09-796-033-2
; Sequence 2, Application US/09796033
; Patent No. US2002000946A1
; GENERAL INFORMATION:
; APPLICANT: Magilavy, Daniel
; TITLE OF INVENTION: METHOD OF MODULATING MEMORY EFFECTOR
; FILE REFERENCE: T-CELLS AND COMPOSITIONS
; FILE REFERENCE: 10274-044001
; CURRENT APPLICATION NUMBER: US/09796.033
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: PCT/US99/20026
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: US 60/098,456
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(28)
US-09-796-033-2

Query Match	100.0%	Score 1326;	DB 10;	Length 250;
Best Local Similarity	100.0%;	Pred. No. 2.6e-112;		
Matches 250;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MVAGSDAGRALGVLSVVVCLLHCFGFCFSQIYGVYGVNVTHTVPSNVPLKEVLWKKQK	60	
Db	1	MVAGSDAGRALGVLSVVVCLLHCFGFCFSQIYGVYGVNVTHTVPSNVPLKEVLWKKQK	60	
QY	61	DKVALENSERAFSPFKNRVYLDTVSGSLTIYNTLTSDEDEYEMESPNTIDTMKFFLYV	120	
Db	61	DKVALENSERAFSPFKNRVYLDTVSGSLTIYNTLTSDEDEYEMESPNTIDTMKFFLYV	120	
QY	121	LESLSPTLTALTNGSIEVQCMIPHEYNHSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND	180	
Db	121	LESLSPTLTALTNGSIEVQCMIPHEYNHSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND	180	


```
;
;
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/730,465
; FILING DATE: 05-Dec-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08755
; FILING DATE: 06-OCT-1992
; APPLICATION NUMBER: US 07/862,022
; FILING DATE: 12-APR-1992
; APPLICATION NUMBER: US 07/770,969
; FILING DATE: 07-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis (PLM)
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGP-111CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-730-465-4

Query Match          94.3%; Score 1250; DB 10; Length 240;
Best Local Similarity 99.6%; Pred. No. 1.8e-105;
Matches 236; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWAGSDAGRALGVLSVVCLLHCFGFISCFSQIYGVVYGNVTFHVPNSVPLKEVLWKQK 60
Db 1 MWAGSDAGRALGVLSVVCLLHCFGFISCFSQIYGVVYGNVTFHVPNSVPLKEVLWKQK 60
QY 61 DKVAELENSEFRAFSSFKNRVYLDTVSGSLTIYNTLTSDEDEYEMESPNITDTMKFFLYV 120
Db 61 DKVAELENSEFRAFSSFKNRVYLDTVSGSLTIYNTLTSDEDEYEMESPNITDTMKFFLYV 120
QY 121 LESPTLTALTNGLSIEVQCMIPSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
Db 121 LESPTLTALTNGLSIEVQCMIPSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
QY 181 LPQKIQCTLSNPLNTSSIIILTCIPSSGHSRHRALPIPLAVITTCIVLYMNGI 237
Db 181 LPQKIQCTLSNPLNTSSIIILTCIPSSGHSRHRALPIPLAVITTCIVLYMNGM 237

RESULT 5
US-10-091-236-17
; Sequence 17, Application US/10091236
; Patent No. US20020168360A1
; GENERAL INFORMATION:
; APPLICANT: DINGIVAN, CHRISTINE A.
; TITLE OF INVENTION: METHODS OF PREVENTING OR TREATING INFLAMMATORY OR AUTOIMMUNE
; DISORDERS BY ADMINISTERING INTEGRIN ALPHA-V-BETA-3 ANTAGONISTS
; TITLE OF INVENTION: COMBINATION WITH OTHER PROPHYLACTIC OR THERAPEUTIC AGENTS
; FILE REFERENCE: 10271-053-999
; CURRENT APPLICATION NUMBER: US/10/091,236
; PRIOR FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: US 60/273,098
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/316,321
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17

Query Match          47.5%; Score 630; DB 10; Length 347;
Best Local Similarity 52.8%; Pred. No. 2.1e-49;
Matches 151; Conservative 22; Mismatches 41; Indels 72; Gaps 12;

QY 1 MWAGSDAGRALGVLSVVCLLHCFGFISCFSQIYGVVYGNVTFHVPNSVPLKEVLWKQK 60
Db 1 MWAGSDAGRALGVLSVVCLLHCFGFISCFSQIYGVVYGNVTFHVPNSVPLKEVLWKQK 60
QY 61 DKVAELENSEFRAFSSFKNRVYLDTVSGSLTIYNTLTSDEDEYEMESPNITDTMKFFLYV 120
Db 61 DKVAELENSEFRAFSSFKNRVYLDTVSGSLTIYNTLTSDEDEYEMESPNITDTMKFFLYV 120
QY 121 LESPTLTALTNGLSIEVQCMIPSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 154
Db 121 DKHTCCPCPAPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPE 172
QY 155 IMYSW-----DCPMEQCKRNST-----SIYFKMENDL--PQKIQCTLSN-----P 192
Db 173 VKFNWYDGVGEVHNAKTKPREE-QYNSTYRVSVTLVHODWLNGLKEYKCKVSNKALPAP 231
QY 193 LFNITSSIIILTCIPSSGHSRHR--YALIP-----IPLAVITTCIV 231
Db 232 TEKTISK-----AKQPREPOVYTLPPSRDELTKNQVSLTCLV 269

RESULT 6
US-09-796-033-8
; Sequence 8, Application US/09796033
; Patent No. US20020009446A1
; GENERAL INFORMATION:
; APPLICANT: Magilavay, Daniel
; TITLE OF INVENTION: METHOD OF MODULATING MEMORY EFFECTOR
; FILE REFERENCE: 10274-044001
; CURRENT APPLICATION NUMBER: US/09/796,033
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: PCT/US99/20026
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: US 60/098,456
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(28)
US-09-796-033-8

Query Match          47.5%; Score 630; DB 10; Length 347;
Best Local Similarity 52.8%; Pred. No. 2.1e-49;
Matches 151; Conservative 22; Mismatches 41; Indels 72; Gaps 12;

QY 1 MWAGSDAGRALGVLSVVCLLHCFGFISCFSQIYGVVYGNVTFHVPNSVPLKEVLWKQK 60
Db 1 MWAGSDAGRALGVLSVVCLLHCFGFISCFSQIYGVVYGNVTFHVPNSVPLKEVLWKQK 60
QY 61 DKVAELENSEFRAFSSFKNRVYLDTVSGSLTIYNTLTSDEDEYEMESPNITDTMKFFLYV 120
Db 61 DKVAELENSEFRAFSSFKNRVYLDTVSGSLTIYNTLTSDEDEYEMESPNITDTMKFFLYV 120
QY 121 LESPTLTALTNGLSIEVQCMIPSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 154
Db 121 DKHTCCPCPAPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPE 172
QY 155 IMYSW-----DCPMEQCKRNST-----SIYFKMENDL--PQKIQCTLSN-----P 192
```



```
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/730,465
; FILING DATE: 05-Dec-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08755
; FILING DATE: 06-OCT-1992
; APPLICATION NUMBER: US 07/862,022
; FILING DATE: 12-APR-1992
; APPLICATION NUMBER: US 07/770,969
; FILING DATE: 07-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis (PLM)
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGP-111CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 351 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-730-465-6
```

```
Query Match      8.7%; Score 115.5; DB 10; Length 351;
Best Local Similarity 23.0%; Pred. No. 0.0049;
Matches 59; Conservative 42; Mismatches 76; Indels 79; Gaps 17;

QY 32 QITGVVGNVTHVPS--NVPLEVLWKKQKD--KVDEL--ENSEFR--AFSFKNRV 81
Db 32 ETWGLAQDINDLIPFQMSDDIDDKWKTSDKKIAQFRKEKTEFKDYLKFN-- 89
QY 82 YLDTVSGSLTYNLSSDEDEY-----ESPNTDTMKFFLYLVLESPLSP----- 127
Db 90 -----GTLKIRHLKTDQDIKVSIDTKGNVLEKI--FDLKIQERVSRKPKISWTGINT 142
QY 128 TLTCALPNSIEVQCMPE-----HYNSHRGLIMYSWDCPMEQCKRNSTSIYFKMEN 179
Db 143 TLTCEVMNGT-----DPELNLYQDGKHLKLSQRVITHKW-----TTS----- 179
QY 180 DLPQKIOCTLSNPLNFTSSIIITCIPSSGHSRHYALIP-----PLAVITTCIVLYM 234
Db 180 -LSAKFKCAGNV-SKESVEPVSC-PEKG--LDIYLIIGTCGGSLMVFVALIVFYI 234
QY 235 NGILKCDKPDRTNSN 250
Db 235 T-----KRRKQSRRN 245
```

```
RESULT 10
US-09-982-107-2
; Sequence 2, Application US/09982107
; Patent No. US20020159958A1
; GENERAL INFORMATION:
; APPLICANT: HIATT, ANDREW C.
; APPLICANT: HEIN, MICHAEL B.
; TITLE OF INVENTION: METHODS FOR PRODUCING IMMUNOGLOBULINS CONTAINING
; TITLE OF INVENTION: PROTECTION PROTEINS IN PLANTS AND THEIR USE
; FILE REFERENCE: EPI3002E
; CURRENT APPLICATION NUMBER: US/09/982,107
; CURRENT FILING DATE: 2001-10-16
```

```
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 773
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-982-107-2
```

```
Query Match      8.0%; Score 105.5; DB 9; Length 773;
Best Local Similarity 22.0%; Pred. No. 0.11;
Matches 47; Conservative 39; Mismatches 79; Indels 49; Gaps 11;

QY 39 GNVTFHP-----SNVPLEVLWKKQDKVAELENSEFRASFKNRVYLDTVSG----S 89
Db 363 GSVTIRCPYNPKRSDSHLQYLWEGSOTRHLVDSGEGVLQKDYTGRLALFEEPGNGTFS 422
QY 90 LTIYNLSSDEDEY---EMESPNTDTMKFFLYLVLESPLSP-----CALTNGSIEVQCM 144
Db 423 VVLNQLTAEDGEGFYWCVSDDDESLLTSVK--LQIVDGEPSPTIDKFTAVQGEPEVITCHF 480
QY 145 P-EHYNSHRGLIMYSWDCPMEQCKRNSTSIYFKMENLDPKIOCTLSNPLNFTSSIIIT 203
Db 481 PCKYFSEK-----YW-----CKWHDG-----EDLPTKL--SSSGDLVKCNLNLVLT 522
QY 204 T-----CIPSSGHSRHYALIP 223
Db 523 LTLDSVSEDDGHWYWCAGKDGHEFEFEVAAVRVEL 556
```

```
RESULT 11
US-09-818-247-6
; Sequence 6, Application US/09818247
; Patent No. US20020102657A1
; GENERAL INFORMATION:
; APPLICANT: Mostov, Keith E.
; APPLICANT: Chapin, Steven J.
; APPLICANT: Richman-Eisenstat, Janice
; TITLE OF INVENTION: Ligands directed to the No. US20020102657A1-Secretory Component
; TITLE OF INVENTION: No. US20020102657A1-stalk Region of p1gR and Methods of Use T
; FILE REFERENCE: 18062E-000910US
; CURRENT APPLICATION NUMBER: US/09/818,247
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: WO PCT/US01/09699
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/192,197
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 60/192,198
; PRIOR FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 773
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
; FEATURE:
; OTHER INFORMATION: rabbit polyclonal immunoglobulin receptor (p1gR)
US-09-818-247-6
```

```
Query Match      8.0%; Score 105.5; DB 10; Length 773;
Best Local Similarity 22.0%; Pred. No. 0.11;
Matches 47; Conservative 39; Mismatches 79; Indels 49; Gaps 11;

QY 39 GNVTFHP-----SNVPLEVLWKKQDKVAELENSEFRASFKNRVYLDTVSG----S 89
Db 363 GSVTIRCPYNPKRSDSHLQYLWEGSOTRHLVDSGEGVLQKDYTGRLALFEEPGNGTFS 422
QY 90 LTIYNLSSDEDEY---EMESPNTDTMKFFLYLVLESPLSP-----CALTNGSIEVQCM 144
Db 423 VVLNQLTAEDGEGFYWCVSDDDESLLTSVK--LQIVDGEPSPTIDKFTAVQGEPEVITCHF 480
QY 145 P-EHYNSHRGLIMYSWDCPMEQCKRNSTSIYFKMENLDPKIOCTLSNPLNFTSSIIIT 203
Db 481 PCKYFSEK-----YW-----CKWHDG-----EDLPTKL--SSSGDLVKCNLNLVLT 522
```

Db 481 PKYFSSK-----YW-----CKNDHGC-----EDLPTKL--SSSGDLVKCNNNMLVLT 522

Qy 204 T-----CIPSSGSHSRHYALIPL 223

Db 523 LTLDSVDEDEGWYWCAGKDGHEFEVAAVRVEL 556

RESULT 12

US-09-732-524-4

; Sequence 4, Application US/09732524

; Patent No. US2002004193A1

; GENERAL INFORMATION:

; APPLICANT: Khodadoust, Mehran

; TITLE OF INVENTION: NOVEL MP-7 PROTEIN AND NUCLEIC ACID MOLECULES

; TITLE OF INVENTION: AND USES THEREOF

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD, LLP

; STREET: 28 State Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/732,524

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/261,759

; FILING DATE:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/090,579

; FILING DATE: 1998-JUN-25

; ATTORNEY/AGENT INFORMATION:

; NAME: Mandragouras, Amy E.

; REGISTRATION NUMBER: 36,207

; REFERENCE/DOCKET NUMBER: MNI-048CP

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617)227-7400

; TELEFAX: (617)742-4214

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 312 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; FRAGMENT TYPE: internal

US-09-732-524-4

Query Match 6.88; Score 90; DB 10; Length 312;

Best Local Similarity 21.74; Pred. No. 0.83;

Matches 53; Conservative 42; Mismatches 107; Indels 42; Gaps 13;

Qy 31 QQIYGVYGVNTHFVPSNV-PLKEVLKKQKDKVAELE-NSEFRAFSEKRNRYLDTVSG 88

Db 4 KELVSGVGAFTPLKSKVKQVDSIVYNTTPTLTIQPEGGTIIVTQNNRERVDFPDG 63

Qy 89 --SLTIYNTSSDEVEY-----EMESPNITDTMKFFLYVLESPLPTLTAL---TNG 136

Db 64 GYSLKSLKKNDGIYVGVYSSSQQFS---TOEYVHVEHLSPKRYTMGLQSNKNG 120

Qy 137 SIEVQ---CMIEPHYNSHRLIMYSWDCPMEQCKRNSTSIYFKM-----ENDLPQKQIT 188

Db 121 TCVNTLTCCM--EHGEED---VIYTKALGQAANESHGSLPISWRWGESDM--TFICV 173

Qy 189 LSPLENTTSSLTITCTI-----PSSGSHSRHYALIPIPLAVITTCIVLYMNGTLKCD 241

Db 174 ARNPVSRNFPSSPILARKLCEGAADDPDSSVLLCLLLVPLLLSLFVLGLFLW---FLKRE 230

Qy 242 RKPD 245

Db 231 RQEE 234

RESULT 13

US-09-992-598-253

; Sequence 253, Application US/09992598

; Patent No. US20020160384A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.

; APPLICANT: Baker, Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Fong, Sherman

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin L.

; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Napier, Mary A.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; FILE REFERENCE: P2730PIC20

; CURRENT APPLICATION NUMBER: US/09/992,598

; CURRENT FILING DATE: 2001-11-14

; PRIOR APPLICATION NUMBER: 60/049787

; PRIOR FILING DATE: 1997-06-16

; PRIOR APPLICATION NUMBER: 60/062250

; PRIOR FILING DATE: 1997-10-17

; PRIOR APPLICATION NUMBER: 60/065186

; PRIOR FILING DATE: 1997-11-12

; PRIOR APPLICATION NUMBER: 60/065311

; PRIOR FILING DATE: 1997-11-13

; PRIOR APPLICATION NUMBER: 60/066770

; PRIOR FILING DATE: 1997-11-24

; PRIOR APPLICATION NUMBER: 60/075945

; PRIOR FILING DATE: 1998-02-25

; PRIOR APPLICATION NUMBER: 60/078910

; PRIOR FILING DATE: 1998-03-20

; PRIOR APPLICATION NUMBER: 60/083322

; PRIOR FILING DATE: 1998-04-28

; PRIOR APPLICATION NUMBER: 60/084600

; PRIOR FILING DATE: 1998-05-07

; PRIOR APPLICATION NUMBER: 60/087106

; PRIOR FILING DATE: 1998-05-28

; PRIOR APPLICATION NUMBER: 60/087607

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/087609

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/087759

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/087827

; PRIOR FILING DATE: 1998-06-03

; PRIOR APPLICATION NUMBER: 60/088021

; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088025

; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088026

;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088028
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088029
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088030
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088033
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088326
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088655
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088734
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088742
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089440
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22

;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 6.8%; Score 90; DB 9; Length 335;
Best Local Similarity 21.7%; Pred. No. 0.91;

Matches 53; Conservative 42; Mismatches 107; Indels 42; Gaps 13;

QY 31 QQIYGVVYGVNTHFVPSNV-PLKVEVLMKKQDKVAELE-NSEFRAFSSFKNRVYLDTVSG 88

Db 27 KELVSGVGAVTFPLKSKVKQVDSIVTFTNTPLVTIQPEGGTIVTQNRNRVDFPDG 86

QY 89 --SUTYNTLTSSDEY-----EMESPNTDTMKFFLYVLESPLSPTLTCL---TNG 136

Db 87 GYSLKSLKKNDSGIYVVGIIYSSSLQOPS---TQEYVLHVYEHLSKPKVTWGLQSNKNG 143
QY 137 STEVQ---CMTEPHYNSHRLIMYNDPCMEOCKENSTSIYFKM-----ENDLPKIOCT 188
Db 144 TCVNLTJTCM--EHGED---VYITWKAIGQAANESHNGSTLPSIWRWGESDM--TFICV 196
QY 189 LSNPLFTNTSIIITTCI-----PSSGHSRHRVALIPIPLAVITTCIVLYMNGILKCD 241
Db 197 ANPVSRRNFSSPILARKLCBGAADDPDSSMWLLCULLVPLLSLFLVLGLFLN---FLKRE 253
QY 242 RRPD 245
Db 254 RQEE 257

RESULT 14
US-09-989-293A-253
; Sequence 253, Application US/09989293A
; Patent No. US20020177164A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C66
; CURRENT APPLICATION NUMBER: US/09/989,293A
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908

Db 254 RQEE 257

Search completed: January 28, 2003, 08:53:00
Job time : 83.077 secs

• • •

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 08:38:45 ; Search time 37.7806 Seconds
(without alignments)
1223.856 Million cell updates/sec

Title: US-09-730-465-8

perfect score: 1856

Sequence: 1 MVAGSDAGRALGVLSVCLL.....MHEALHNHYTKSLSLSPGK 347

Scoring table: BLOSUM62

Learning curve: 2000000
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs. 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Minimum	DB seq	length:
Maximum	DB seq	length: 2000000000

Post-processing: Minimum Match 0%

Post-processing: Minimum Match 0%
Maximum Match 100%

Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_101002:*

1:	/SID52/cgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SID52/cgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SID52/cgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SID52/cgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SID52/cgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SID52/cgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SID52/cgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SID52/cgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SID52/cgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SID52/cgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SID52/cgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SID52/cgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SID52/cgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SID52/cgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SID52/cgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SID52/cgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SID52/cgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SID52/cgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SID52/cgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SID52/cgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SID52/cgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SID52/cgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SID52/cgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES							
Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	1856	100.0	347	13	AAR27163		CD2 binding LFA-3-
2	1856	100.0	347	21	AAV83136		Human transmembran
3	1856	100.0	347	22	AAB61160		Human LFA3TIP fusi
4	1856	100.0	347	22	AAU76228		LFA3TIP fusi
5	1834	98.8	347	14	AAR34224		LFA3TIP fusi
6	1265	68.2	458	22	AAV97596		Amino acid sequenc
7	1263.5	68.1	458	22	AAV97594		Fc1 receptor fusi
8	1262	68.0	744	21	AAV94408		Fc1 receptor fusi
9	1255.5	67.6	633	21	AAV84965		Human VCAM-1/IgG1-
10	1255	67.6	477	20	AAW90207		Anhydro acid sequenc
							Hb γ .2Fc soluble fu

```
XX 12-MAR-1991; 91US-0667971.
PR 07-OCT-1991; 91US-0770967.
XX (BIOJ ) BIOGEN INC.
XX Miller GT, Rosa MD, Wallner BP;
XX WPI; 1992-309760/38.
DR N-PSDB; AAQ28684.
XX
PT CD2-binding domain of lymphocyte function associated antigen-3
PT and DNA - for diagnosing and treating inflammation and
PT auto-immune diseases, e.g. systemic lupus erythematosus and
PT rheumatoid arthritis
XX
PS Claim 13; Fig 12; 85pp; English.
XX
CC The plasmid pSAB152 contains the DNA sequence encoding the LFA-3
CC signal sequence, the amino terminal 92 amino acids of mature LFA-3,
CC ten amino acids of the hinge region of IgG1 and the CH2 and CH3
CC constant domains of IgG1 (see AAQ28678-9 and AAQ28681-2 for details of
CC the construction of pSAB152). A NotI fragment containing the coding
CC sequence of pSAB152 was used in the construction of expression
CC vector pMDR(92)Ig-3 which can be stably maintained in CHO cells to
CC achieve continuous expression of LFA3TIP. The fusion protein can
CC bind to CD2 and inhibit T cell activation, making it useful to
CC treat acute and chronic inflammation, autoimmune disease and
CC in immunomodulation.
XX
SQ Sequence 347 AA;
Query Match 100.0%; Score 1856; DB 13; Length 347;
Best Local Similarity 100.0%; Pred. No. 2.1e-128;
Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVAGSDAGRALGVLSVCLLHCFGFTSCFSQOIYGVVYGNVTFHVPNSVPLKEVLKKQK 60
DB 1 MVAGSDAGRALGVLSVCLLHCFGFTSCFSQOIYGVVYGNVTFHVPNSVPLKEVLKKQK 60
QY 61 DKVAELENSEFRASFSSFNKRVYLDTVSGSLTIYNLTSSDEDEYEMESPNITDTMKFFLV 120
DB 61 DKVAELENSEFRASFSSFNKRVYLDTVSGSLTIYNLTSSDEDEYEMESPNITDTMKFFLV 120
QY 121 DKHTHTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYD 180
DB 121 DKHTHTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYD 180
QY 181 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 240
DB 181 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 240
QY 241 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 300
DB 241 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 300
QY 301 DGSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTQKSLSLSPGK 347
DB 301 DGSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTQKSLSLSPGK 347
RESULT 2
AAV83136
ID AAY83136 standard; Protein; 347 AA.
XX
AC AAY83136;
XX
XX
DT 24-JUL-2000 (first entry)
XX
XX Human transmembrane LFA-3/IgG fusion protein LFA3TIP.
XX
KW LFA3; CD2; cell signalling; modulation; lymphocyte; T cell;
KW memory effector T lymphocyte; psoriatic arthritis;
```

```
KW rheumatoid arthritis; multiple sclerosis; atopic dermatitis;
KW uveitis; inflammatory bowel disease; Crohn's disease;
KW ulcerative colitis; cutaneous T cell lymphoma; inhibition; treatment;
KW therapy.
XX
OS Homo sapiens.
XX
PN WO200012113-A2.
XX
PD 09-MAR-2000.
XX
PF 31-AUG-1999; 99WO-US20026.
XX
PR 31-AUG-1998; 98US-0098456.
XX (BIOJ ) BIOGEN INC.
XX
XX Magilavay D;
XX
XX WPI; 2000-282928/24.
DR N-PSDB; AAZ93401.
XX
PT Selective modulation of memory effector T lymphocytes by administration
PT of a CD2 binding agent which inhibits the CD2/LFA-3 interaction useful
PT for treating conditions such as inflammatory bowel diseases, psoriatic
PT arthritis
XX
PS Disclosure; Page 75-76; 76pp; English.
XX
CC Modulation of LFA3/CD2 interaction by administration of a CD2 binding
CC agent inhibits CD2 signalling and T cell proliferation and activation
CC and more particularly modulates the number and/or distribution of
CC memory effector T lymphocytes. The method can be used for treating a
CC condition in a subject where the condition is characterized by memory
CC effector T lymphocytes playing a role in the pathogenesis of the
CC condition such as psoriatic arthritis, rheumatoid arthritis,
CC multiple sclerosis, atopic dermatitis, uveitis, inflammatory bowel
CC disease, Crohn's disease, ulcerative colitis and cutaneous T cell
CC lymphoma and where the method comprises administering to the subject
CC an amount of CD2 binding agent sufficient to modulate the memory
CC effector T lymphocytes. The methods provide inhibition of antigen
CC specific interactions for all antigens present, inhibition of T cell
CC activation, no general immunosuppression, and possibly induction of
CC tolerance.
XX
SQ Sequence 347 AA;
Query Match 100.0%; Score 1856; DB 21; Length 347;
Best Local Similarity 100.0%; Pred. No. 2.1e-128;
Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVAGSDAGRALGVLSVCLLHCFGFTSCFSQOIYGVVYGNVTFHVPNSVPLKEVLKKQK 60
DB 1 MVAGSDAGRALGVLSVCLLHCFGFTSCFSQOIYGVVYGNVTFHVPNSVPLKEVLKKQK 60
QY 61 DKVAELENSEFRASFSSFNKRVYLDTVSGSLTIYNLTSSDEDEYEMESPNITDTMKFFLV 120
DB 61 DKVAELENSEFRASFSSFNKRVYLDTVSGSLTIYNLTSSDEDEYEMESPNITDTMKFFLV 120
QY 121 DKHTHTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYD 180
DB 121 DKHTHTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYD 180
QY 181 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 240
DB 181 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 240
QY 241 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 300
DB 241 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 300
QY 301 DGSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTQKSLSLSPGK 347
DB 301 DGSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTQKSLSLSPGK 347
```

Db 301 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPGK 347

RESULT 3
AAB61160
ID AAB61160 standard; Protein; 347 AA.

XX AC AAB61160;

XX DT 30-MAR-2001 (first entry)

XX DE Human LFA3TIP fusion protein.

XX KW Human; LFA3TIP; antipsoriatic; dermatological; antiinflammatory;
KW antiallergic; vasotropic; skin condition; T-cell adhesion molecule CD2;
KW lymphocyte function-associated antigen-3; LFA-3; atopic dermatitis;
KW cutaneous T cell lymphoma; fusion.

XX OS Homo sapiens.

XX PN US6162432-A.

XX PD 19-DEC-2000.

XX PF 06-JUN-1995; 95US-0466465.

XX PR 07-OCT-1991; 91US-0770969.

XX PR 02-APR-1992; 92US-0862022.

XX PR 06-OCT-1992; 92WO-US08755.

XX PA (UNMI) UNIV MICHIGAN.

XX PA (BIOJ) BIOGEN INC.

XX PI Cooper KD, Wallner BP;

XX DR WPT; 2001-122130/13.

XX PT Prophylaxis or treatment of skin conditions such as dermatitis,
PT alopecia areata, psoriasis, urticaria in mammals, by administering
PT inhibitor of T-cell adhesion molecule CD2/lymphocyte
PT function-associated antigen-3 -

XX PS Claim 11; Column 39-42; 23pp; English.

XX CC The present sequence is given in a specification relating to a method
CC for treating or providing prophylaxis for a skin condition characterised
CC by increased T cell activation and abnormal antigen presentation in the
CC dermis and epidermis. The method comprises administering to a
CC mammal an inhibitor of T-cell adhesion molecule CD2/lymphocyte
CC function-associated antigen-3 (LFA-3) interaction selected from a CD2
CC polypeptide, an LFA-3 polypeptide, an anti-CD2 and an anti-LFA-3 antibody
CC homologue. The method is useful for treating or providing prophylaxis for
CC a skin condition such as atopic dermatitis, cutaneous T cell lymphoma
CC such as mycosis fungoides, allergic and irritant contact dermatitis,
CC lichen planus, alopecia areata, pyoderma gangrenosum, vitiligo, ocular
CC cicatricial pemphigoid, psoriasis and urticaria or skin condition
CC associated with UV damage. The method results in more specific targeting
CC of therapy to T cells in the initiating stage of the lesion without any
CC effect on polymorphonuclear leukocytes or macrophage mediated effector
CC mechanisms. The patients are less susceptible to infections than with
CC steroids or other general immunosuppressants. The present sequence
CC is a fusion protein comprised of the first extracellular domain of
CC LFA-3 fused to the hinge, CH2 and CH3 regions of human IgG1.

XX SQ Sequence 347 AA;

Query Match 100.0%; Score 1856; DB 22; Length 347;
Best Local Similarity 100.0%; Pred. No. 2.1e-128;
Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MWAGSDAGRALGVSVVCLLHCFGIFSCFSQIYGVVYGNVTFHVPNSVPLKEVLWKQK 60
|||||

Db 1 MWAGSDAGRALGVSVVCLLHCFGIFSCFSQIYGVVYGNVTFHVPNSVPLKEVLWKQK 60
|||||

Qy 61 DKVAELENSERAFSFKNRVYLDTVSGSLTIYNLTSDEDEYEMESPNITDTMKFFLYV 120
|||||
Db 61 DKVAELENSERAFSFKNRVYLDTVSGSLTIYNLTSDEDEYEMESPNITDTMKFFLYV 120
|||||
Qy 121 DKTHTCPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 180
|||||
Db 121 DKTHTCPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 180
|||||
Qy 181 GVEVHNAKTPREEQYNSTRYRVSVLTVLHQDLNGKEYCKVSNKALPAPIEKTISKAK 240
|||||
Db 181 GVEVHNAKTPREEQYNSTRYRVSVLTVLHQDLNGKEYCKVSNKALPAPIEKTISKAK 240
|||||
Qy 241 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 300
|||||
Db 241 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 300
|||||
Qy 301 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPGK 347
|||||
Db 301 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPGK 347
|||||

RESULT 4

AAU76228

ID AAU76228 standard; Protein; 347 AA.

XX AC AAU76228;

XX DT 08-MAY-2002 (first entry)

XX DE LFA3/IgG fusion protein (LFA3TIP).

XX KW LFA-3; lymphocyte function-associated antigen; CD2; dermatological;
KW cytostatic; antiallergic; antiinflammatory; ophthalmic; dermis;
KW epidermis; ski disorder; lymphoma; mycosis fungoides; lichen planus;
KW allergic dermatitis; irritant contact dermatitis; alopecia areata;
KW pyoderma gangrenosum; vitiligo; ocular cicatricial pemphigoid;
KW psoriasis; UV damage; urticaria; LFA3TIP.

XX OS Mammalia.

XX PH Key Location/Qualifiers

FT Peptide 1..28

FT Protein /note= "Signal peptide"

FT Protein 29..347 /note= "Mature LFA3TIP protein"

XX US2002009449-A1.

XX PD 24-JAN-2002.

XX PF 05-DEC-2000; 2000US-0730465.

XX PR 07-OCT-1991; 91US-0770969.

XX PR 02-APR-1992; 92US-0862022.

XX PR 06-OCT-1992; 92WO-US08755.

XX PR 06-JUN-1995; 95US-0466465.

XX PA (BIOJ) BIOGEN INC.

XX PI Wallner BP, Cooper KD;

XX DR WPT; 2002-179118/23.

XX DR N-PSDB; ABK15720.

XX PT Treating conditions characterised by increased T cell activation and
PT abnormal antigen presentation in the dermis and epidermis, e.g. atopic
PT dermatitis, vitiligo and UV damage, by administering an inhibitor of
PT the CD2/LFA-3 interaction -

XX PS Example 3; Page 21-22; 32pp; English.

XX CC This invention relates to a novel method for preventing or treating skin

CC conditions characterised by increased T cell activation and abnormal
CC antigen presentation in the dermis and epidermis. The method comprises
CC inhibition of the CD2/lymphocyte function associated antigen-3 (LFA-3)
CC interaction by administering to a mammal, including a human, an
CC inhibitor of the CD2/LFA-3 interaction. These inhibitors may be anti-
CC LFA-3 antibody homologues, anti-CD2 antibody homologues, soluble LFA-3
CC peptides, soluble CD-2 polypeptides, small molecules such as
CC carbohydrates, LFA-3 and CD2 mimetic agents. This method is useful to
CC treat skin disorders such as cell lymphoma, mycosis fungoides, allergic
CC and irritant contact dermatitis, lichen planus, alopecia areata,
CC pyoderma gangrenosum, vitiligo, ocular cicatricial pemphigoid,
CC psoriasis, UV damage and urticaria. The present sequence represents the
CC mammalian LFA3/Ig fusion protein (LFA3TIP) sequence used to create the
CC inhibitors of CD-2/LFA interaction of the invention.
XX
SQ Sequence 347 AA;

Query Match 100.0%; Score 1856; DB 23; Length 347;
Best Local Similarity 100.0%; Pred. No. 2.1e-128;
Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVAGSDAGRALGVLSVWCLLHCFGFISCFSSQIYGVYGVNVTFFHVPNSVPLKEVLWKKQK 60
Db 1 MVAGSDAGRALGVLSVWCLLHCFGFISCFSSQIYGVYGVNVTFFHVPNSVPLKEVLWKKQK 60
Qy 61 DKVAELENSEFRAFSSFKNRVYLDTVSGSLTIYNTLTSSDEDEYEMESPNITDTMKFFLYV 120
Db 61 DKVAELENSEFRAFSSFKNRVYLDTVSGSLTIYNTLTSSDEDEYEMESPNITDTMKFFLYV 120
Qy 121 DKHTCPCPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 180
Db 121 DKHTCPCPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 180
Qy 181 GVEVHNNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 240
Db 181 GVEVHNNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 240
Qy 241 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 300
Db 241 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 300
Qy 301 DGSFFLYSKLTVDKSRWQOGNVFSCSVMHREALHNYTQKLSLSLSPGK 347
Db 301 DGSFFLYSKLTVDKSRWQOGNVFSCSVMHREALHNYTQKLSLSLSPGK 347

RESULT 5
AAR34224
ID AAR34224 standard; Protein: 347 AA.

XX AAR34224;
XX
XX 09-AUG-1993 (first entry)
XX Amino acid sequence of LFA3TIP.
XX
XX LFA-3; CD2 binding protein; graft; xenograft; allograft;
XX fusion protein.
XX
XX Synthetic.
XX
XX WO9306852-A.
XX
XX 15-APR-1993.
XX
XX 06-OCT-1992; 92WO-US08754.
XX
XX 07-OCT-1991; 91US-0772705.
XX 12-MAR-1992; 92US-0850706.
XX
XX (BIOJ) BIOGEN INC.
XX
XX Benjamin CD, Wallner BP;

XX WPI: 1993-134134/16.
DR N-PSDB: AAQ40423.
XX
PT Improving tolerance of transplanted allo:graft or xenograft
PT tissue - using LFA-3 or CD-2 binding protein esp. for humans
PT given cardiac or renal grafts
XX
PS Disclosure; Page 58-59; 68pp; English.
XX
CC A plasmid, pSAB152, encoding LFA3TIP is deposited under ATCC 68720.
CC The fusion protein comprises the amino terminal 92 amino acids of
CC mature LFA-3, the C-terminal 10 amino acids of a human IgG1 hinge
CC region containing the two cysteine residues thought to participate
CC in interchain disulfide bonding, and the CH2 and CH3 regions of a
CC human IgG1 heavy chain constant domain. One way of producing
CC LFA3TIP is described in United States patent application
CC 07/770,967.
XX

SQ Sequence 347 AA;

Query Match 98.8%; Score 1834; DB 14; Length 347;
Best Local Similarity 98.8%; Pred. No. 8.8e-127;
Matches 343; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MVAGSDAGRALGVLSVWCLLHCFGFISCFSSQIYGVYGVNVTFFHVPNSVPLKEVLWKKQK 60
Db 1 MVAGSDAGRALGVLSVWCLLHCFGFISCFSSQIYGVYGVNVTFFHVPNSVPLKEVLWKKQK 60
Qy 61 DKVAELENSEFRAFSSFKNRVYLDTVSGSLTIYNTLTSSDEDEYEMESPNITDTMKFFLYV 120
Db 61 DKVAELENSEFRAFSSFKNRVYLDTVSGSLTIYNTLTSSDEDEYEMESPNITDTMKFFLYV 120
Qy 121 DKHTCPCPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 180
Db 121 DKHTCPCPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 180
Qy 181 GVEVHNNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 240
Db 181 GVEVHNNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 240
Qy 241 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 300
Db 241 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 300
Qy 301 DGSFFLYSKLTVDKSRWQOGNVFSCSVMHREALHNYTQKLSLSLSPGK 347
Db 301 DGSFFLYSKLTVDKSRWQOGNVFSCSVMHREALHNYTQKLSLSLSPGK 347

RESULT 6
AA97596
ID AA97596 standard; Protein: 458 AA.

XX AA97596;
XX
XX 05-APR-2001 (first entry)
XX
XX Flt1 receptor fusion protein VEGFR1R2-FcDeltaC1(a).
XX
XX Flt1 receptor; fusion protein; chimeric protein; pharmacokinetic;
XX plasma leakage; vascular permeability; IgG Fc region.
XX
XX Unidentified.
XX
XX WO200075319-A1.
XX
XX 14-DEC-2000.
XX
XX 23-MAY-2000; 2000WO-US14142.
XX
XX 08-JUN-1999; 99US-0138133.
XX

PA (REGE-) REGENERON PHARM INC.
XX Papadopoulos NJ, Davis S, Yancopoulos GD;
PI WPI; 2001-071076/08.
DR N-PSDB; AAA91076.
XX Nucleic acid molecule encoding mammalian phospholipid transfer protein,
PT and its fragments, useful for diagnosis, evaluation, and treatment of
PT diseases associated with the gene expression and for producing model
PT systems -
XX Claim 49; Fig 24; 159pp; English.
XX This sequence represents a fusion protein of the invention between
CC the Flt1 receptor and the Fc region of IgG. The specification relates
CC to modified chimeric polypeptides with improved pharmacokinetics. The
CC modified chimeric polypeptides are preferably Flt1 receptor polypeptides
CC that have been modified to improve their pharmacokinetic profile. The
CC polypeptides can be used to decrease or inhibit plasma leakage and/or
CC vascular permeability in a mammal.
XX Sequence 458 AA;
SQ
Query Match 68.2%; Score 1265; DB 22; Length 458;
Best Local Similarity 78.7%; Pred. No. 7.7e-85;
Matches 248; Conservative 15; Mismatches 40; Indels 12; Gaps 4;
QY 40 NVTFHVPSPNPLKEVLWK--KQDKVAELENSEFRASFKNRYVLDTVSGSLTIYNLTS 97
Db 149 NCTARTELNVGI-DFNWEYPSKQHKKLVNRDLKTQSGSEMKKFLST----LTIDGVTR 203
QY 98 SDEDEYEMESPNTIDTMKFLYV-----DKTHTCPCPAPPELLGGPSVFLFPPKDTLM 152
Db 204 SDGLYTCAASSGLMTKKNSTFVRVHEKDKTHTCPCPAPPELLGGPSVFLFPPKDTLM 263
QY 153 ISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKRPREQNSYIRVSVLTCLV 212
Db 264 ISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKRPREQNSYIRVSVLTCLV 323
QY 213 WLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYITLPPSRDELTKNQVSLTCLVKGF 272
Db 324 WLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYITLPPSRDELTKNQVSLTCLVKGF 383
QY 273 YPSDIAEVESNGQPENNYKTTPPVLDSDGSFFLYSLKLTVDKSRWQQGNVFCSV 332
Db 384 YPSDIAEVESNGQPENNYKTTPPVLDSDGSFFLYSLKLTVDKSRWQQGNVFCSV 443
QY 333 HNHYTQKSLSPGK 347
Db 444 HNHYTQKSLSPGK 458
RESULT 7
AA97594
ID AAY97594 standard; Protein: 458 AA.
XX AAY97594;
AC AAY97594;
XX 05-APR-2001 (first entry)
XX Flt1 receptor fusion protein Flt1D2.Flt1D3.FcDeltaC1(a).
XX Flt1 receptor; fusion protein; chimeric protein; pharmacokinetic;
KW plasma leakage; vascular permeability; IgG Fc region.
XX Unidentified.
OS
XX WO200075319-A1.
PN
XX 14-DEC-2000.
PD
XX 23-MAY-2000; 2000WO-US14142.
PF

XX 08-JUN-1999; 99US-0138133.
PR (REGE-) REGENERON PHARM INC.
XX Papadopoulos NJ, Davis S, Yancopoulos GD;
PI WPI; 2001-071076/08.
DR N-PSDB; AAA91074.
XX Nucleic acid molecule encoding mammalian phospholipid transfer protein,
PT and its fragments, useful for diagnosis, evaluation, and treatment of
PT diseases associated with the gene expression and for producing model
PT systems -
XX Claim 49; Fig 21; 159pp; English.
XX This sequence represents a fusion protein of the invention between
CC the Flt1 receptor and the Fc region of IgG. The specification relates
CC to modified chimeric polypeptides with improved pharmacokinetics. The
CC modified chimeric polypeptides are preferably Flt1 receptor polypeptides
CC that have been modified to improve their pharmacokinetic profile. The
CC polypeptides can be used to decrease or inhibit plasma leakage and/or
CC vascular permeability in a mammal.
XX Sequence 458 AA;
SQ
Query Match 68.1%; Score 1263.5; DB 22; Length 458;
Best Local Similarity 78.0%; Pred. No. 9.9e-85;
Matches 248; Conservative 15; Mismatches 40; Indels 15; Gaps 4;
QY 40 NVTFHVPSPNPLKEVLWK--KQDKVAELENSEFRASFKNRYVLDTVSGSLTIYNLTS 97
Db 146 NCTARTELNVGI-DFNWEYPSKQHKKLVNRDLKTQSGSEMKKFLST----LTIDGVTR 200
QY 98 SDEDEYEMESPNTIDTMKFLYV-----DKTHTCPCPAPPELLGGPSVFLFPPKPD 149
Db 201 SDGLYTCAASSGLMTKKNSTFVRVHEKDGKTHTCPCPAPPELLGGPSVFLFPPKPD 260
QY 150 TLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKRPREQNSYIRVSVLTCLV 209
Db 261 TLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKRPREQNSYIRVSVLTCLV 320
QY 210 HDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYITLPPSRDELTKNQVSLTCLV 269
Db 321 HDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYITLPPSRDELTKNQVSLTCLV 380
QY 270 KGFYPSDIAEVESNGQPENNYKTTPPVLDSDGSFFLYSLKLTVDKSRWQQGNVFCSV 329
Db 381 KGFYPSDIAEVESNGQPENNYKTTPPVLDSDGSFFLYSLKLTVDKSRWQQGNVFCSV 440
QY 330 EALHNHYTQKSLSPGK 347
Db 441 EALHNHYTQKSLSPGK 458
RESULT 8
AAY94408
ID AAY94408 standard; Protein: 744 AA.
XX AAY94408;
AC AAY94408;
XX 11-SEP-2000 (first entry)
XX Human VCAM-1/IgG1-Fc fusion protein.
XX Human; cellular adhesion molecule; ACAM; nontropic; antiepileptic;
KW neuroleptic; renal-active; antidiabetic; neuroactive; neuroprotectant;
KW dementia; epilepsy; schizophrenia; peripheral nerve injury;
KW diabetic neuropathy; fusion protein.
XX Homo sapiens.
OS
XX

PN WC200032633-AL.
 XX 08-JUN-2000.
 XX 02-DEC-1999; 99WO-US28878.
 XX 02-DEC-1998; 98US-0203462.
 XX (ICOS-) ICOS CORP.
 XX Hoekstra DM, Loughney K, Stauton DE, Vazeux R;
 XX WPI: 2000-422952/36.
 DR N-PSDB; AAA30442.
 XX Nucleic acids encoding ACAM, a human cellular adhesion molecule, useful
 PT for diagnosing, preventing and treating diseases associated with ACAM
 PT expression and activity, e.g. epilepsy and schizophrenia -
 XX Example 5; Page 174-176; 187pp; English.
 XX The present sequence is a fusion protein comprising part of human
 CC VCAM-1 and the Fc region of human IgG1. The sequence encoding this
 CC protein was generated by integrating the VCAM-1 coding region into a
 CC plasmid designated pDEF24/IgG1, which encodes the hinge and constant
 CC CH2-CH3 domains of human IgG1. The sequence was then ligated
 CC to the expression vector pDEF24 and transformed into DH5a competent
 CC cells. The fusion protein was used as a control in adhesion experiments
 CC involving ACAM#4/IgG-Fc and ACAM#5/IgG4-Fc fusion proteins. ACAM#4 and
 CC ACAM#6 are human foetal brain cDNA clones containing alternatively
 CC spliced full-length variants of a novel adhesion molecule designated
 CC ACAM. ACAM nucleic acids and polypeptides may be used in the prevention,
 CC treatment and diagnosis of diseases associated with inappropriate ACAM
 CC expression and activity such as dementia, epilepsy, schizophrenia,
 CC peripheral nerve injuries and diabetic neuropathies. They may be used to
 CC rectify mutations or deletions in a patient's genome that affect the
 CC activity of ACAM or to supplement insufficient ACAM production in a
 CC patient. Conversely, antisense nucleic acid molecules may be
 CC administered to down-regulate ACAM expression. The nucleotide sequence
 CC may also be used as a DNA probe in diagnostic assays (e.g. PCR) to
 CC detect and quantitate the presence of similar nucleic acid sequences in
 CC samples, and hence determine which patients may be in need of
 CC restorative therapy. ACAM polypeptides may be used as antigens in the
 CC production of antibodies against ACAM and in assays to identify
 CC modulators (agonists and antagonists) of ACAM expression and activity.
 XX Sequence 744 AA;
 SQ

Query Watch 68.0%; Score 1262; DB 21; Length 744;
 Best Local Similarity 75.3%; Pred. NO. 2.3e-84;
 Matches 253; Conservative 15; Mismatches 40; Indels 28; Gaps 6;
 QY 35 GVYGN-...VTFHVPSPVPLKEVLWKQKDKVAELENSEPRFSSPKN-----RVYLD 84
 Db 414 GLVNGSSVTSCKVPYVPLDLLEILLKGETI-LENIEFLEDTKSLNKSLEMTIP 472
 QY 85 TVSG---SLTIYNTLSDEDEYEMESPNTDTMKFFLYV-----DKTHTCCPCPA 131
 Db 473 TIEDTGKALVCQAKLHDDMEFEFKQKSTQT---LYVNVAPVDPKSCDKTHTCCPCPA 528
 QY 132 PELLGGPSVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKP 191
 Db 529 PELLGGPSVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKP 588
 QY 192 REEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIETISKAGQPRPQVYTL 251
 Db 589 REEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIETISKAGQPRPQVYTL 648
 QY 252 PPSRDLTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLDSDGSFFLYSKLT 311
 Db 649 PPSRDLTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLDSDGSFFLYSKLT 708
 QY 312 VDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 347

Db 709 VDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 744
 RESULT 9
 AAY84965
 ID AAY84965 standard; Protein; 633 AA.
 XX AC AAY84965;
 XX 21-AUG-2000 (first entry)
 XX Amino acid sequence of a CD-20 specific chimeric receptor.
 KW CD20-specific receptor; CD-20 specific redirected T cell; leukemia;
 KW CD20+ malignancy; non-Hodgkin's lymphoma; myeloablative chemotherapy;
 KW stem cell rescue; autoimmune disease; lupus; rheumatoid arthritis.
 XX Synthetic.
 OS Mus sp.
 XX Key Location/Qualifiers
 FH Peptide 27..86
 FT /note= "signal peptide from murine T86.66 antibody
 FT kappa light chain"
 FT Region 21..126
 FT /note= "anti-CD20 variable regions"
 FT Peptide 145..266
 FT /note= "GS18 linker"
 FT Region 283..392
 FT /note= "hinge region"
 FT Region 393..499
 FT /note= "CH3 region"
 FT Region 500..521
 FT /note= "CD4 transmembrane region"
 FT Region 522..633
 FT /note= "zeta chain"
 XX WO200023573-A2.
 XX 27-APR-2000.
 XX 20-OCT-1999; 99WO-US24484.
 XX 20-OCT-1998; 98US-0105014.
 XX (CITY) CITY OF HOPE.
 XX Raubitschek A, Jensen MC, Wu AM;
 DR WPI: 2000-339676/29.
 DR N-PSDB; AAA15019.
 XX Genetically engineered CD20-specific redirected T cells useful for
 PT treating a CD20+ malignancy, such as non-Hodgkin's lymphoma or CD20+
 PT acute or chronic leukemia, and autoimmune disease -
 XX Claim 10; Page 53-55; 58pp; English.
 XX The present sequence represents a synthetic CD20-specific chimeric
 CC receptor. The specification describes CD-20 specific redirected T cells
 CC which express and bear on the cell surface membrane a CD20-chimeric
 CC receptor comprising an intracellular signalling domain, a transmembrane
 CC domain and an extracellular domain, the extracellular domain comprising
 CC a CD20-specific receptor. The genetically engineered CD20-specific
 CC redirected T cells are useful for treating a CD20+ malignancy, such
 CC as non-Hodgkin's lymphoma or CD20+ acute or chronic leukemia, in a
 CC human patient having previously undergone myeloablative chemotherapy and
 CC stem cell rescue. The genetically engineered CD20-specific redirected
 CC T cells are also useful for abrogating an untoward B cell function, such
 CC as autoimmune disease (lupus or rheumatoid arthritis) in a patient.
 XX Sequence 633 AA;
 SQ

Query Match 67.6%; Score 1255.5; DB 21; Length 633;
Best Local Similarity 82.5%; Pred. No. 5.6e-84;
Matches 241; Conservative 5; Mismatches 25; Indels 21; Gaps 2;

QY 77 FKNRVLDVTVSGSLTIY-----NLTSSDEDEYEMESPNTIDTMKFFLVV----- 120
DB 208 FKGKATLTADKSSSTAYMQLSSLTSEDSADYYCARSNYYGSSYWFEDVWAGATTIVVSSL 267
QY 121 -----DKTHTCTPCPAPPELLGGPSVFLFPPKPKDTLMISRPETVCVTVVVDVSHEDPEVKF 175
DB 268 DKSSDKTHTCTPCPAPPELLGGPSVFLFPPKPKDTLMISRPETVCVTVVVDVSHEDPEVKF 327
QY 176 NWTVDGVEVHNAKTPREEQVNSTYRVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 235
DB 328 NWTVDGVEVHNAKTPREEQVNSTYRVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 387
QY 236 ISKAKQCPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTP 295
DB 388 ISKAKQCPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTP 447
QY 296 PVLDSGSPFLYSKLTVDKSRWQGNVFCFSVMHEALHNHYTQKSLSLSPCK 347
DB 448 PVLDSGSPFLYSKLTVDKSRWQGNVFCFSVMHEALHNHYTQKSLSLSPCK 499

RESULT 10
AAW90207
ID AAW90207 standard; Protein; 477 AA.
AC AAW90207;
XX 10-MAY-1999 (first entry)
DT hb7.2Fc soluble fusion protein.
DE
XX B7 binding molecule; costimulatory molecule; B7.1; CD80; B7.2;
KW CD86; T cell activation; inhibitor; graft versus host disease;
KW transplant rejection; allograft rejection; autoimmune disease;
KW allergy; therapy; human; antibody; hb7.1fc.
XX Chimeric - Homo sapiens.
OS Chimeric - synthetic.

XX Key Location/Qualifiers
FT Peptide 1..16 /note= "potential eukaryotic secretory signal"
FT peptide
FT Domain 17..239 /note= "human B7.2 (mature protein) extracellular
FT domain"
FT Peptide 240..245 /note= "introduced by PCR cloning strategy"
FT Protein 246..477 /note= "human IgG1-Fc (hinge-CH2-CH3)"
FT
FT
XX WO9858965-A2.
PN
XX
XX 30-DEC-1998.
PD
XX 22-JUN-1998; 98WO-EP03791.
PF
XX 20-JUN-1997; 97EP-0870092.
PR
XX (INNO-) INNOGENETICS NV.
PA
XX Bosman A, Buyse M, Lorre K, Sablon E;
PI
XX WPT; 1999-105615/09.
DR
XX New molecules which bind B7.1 and B7.2 - useful to prevent and treat
PT immune diseases including allograft rejection
PR
XX

PS Example 3.1.1.3; Fig 3; 182pp; English.
XX This 54 kDa soluble fusion protein, termed hb7.2Fc, is composed of
CC human co-stimulatory molecule B7.2 extracellular domain fused
CC C-terminally to human IgG1-Fc. It was produced by PCR
CC amplification of hb7.2 cDNA in plasmid pcDNAIneoHB7.2, and
CC insertion of the amplified cDNA into pVL-Fc (ICCG3048), resulting
CC in pVishb7.2-Fc (ICCG3004) baculotransfer plasmid. The invention
CC relates to molecules such as diabodies, trivalent and tetraivalent
CC antibodies and small antigen binding peptides which can cross-link,
CC or cross-react with, B7.1 and B7.2 expressed on professional
CC antigen-presenting cells leading to the inhibition of
CC antigen-specific T cell activation. Methods to produce such
CC molecules are provided. The molecules are used to treat or prevent
CC diseases of the immune system, in particular graft rejection, graft
CC versus host disease, allergy and autoimmune diseases (claimed).
XX

SQ Sequence 477 AA;
Query Match 67.6%; Score 1255; DB 20; Length 477;
Best Local Similarity 78.3%; Pred. No. 4.4e-84;
Matches 242; Conservative 15; Mismatches 26; Indels 26; Gaps 4;

QY 55 LMKKQKDKVAELENSSEFRAFSGFKNRVLDVTVSGSLTYNLTSSDED-----EYEME- 106
DB 179 IMQKSQDNVTLEYDVSLSVSFDP-----VTSNMTIFCILETDKTRLLSSPFSIELED 232
QY 107 -----SPNITDTMKFFLYVDKTHCTPCPAPPELLGGPSVFLFPPKPKDTLMISRTPE 158
DB 233 FQPPPDHRSFGLQEPKS----CDKTHCTPCPAPPELLGGPSVFLFPPKPKDTLMISRTPE 288
QY 159 VTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQVNSTYRVSVLTVLHQDWLNGKE 218
DB 289 VTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQVNSTYRVSVLTVLHQDWLNGKE 348
QY 219 YKCKVSNKALPAPIEKTISKAKGQPREPVYTLPPSRDELTKNQVSLTCLVKGFYPSDIA 278
DB 349 YKCKVSNKALPAPIEKTISKAKGQPREPVYTLPPSRDELTKNQVSLTCLVKGFYPSDIA 408
QY 279 VEWESNGQPENNYKTTTPVLDSDGSPFLYSKLTVDKSRWQGNVFCFSVMHEALHNHYTQ 338
DB 409 VEWESNGQPENNYKTTTPVLDSDGSPFLYSKLTVDKSRWQGNVFCFSVMHEALHNHYTQ 468
QY 339 KSLSLSPCK 347
DB 469 KSLSLSPCK 477

RESULT 11
AAW73514
ID AAW73514 standard; Protein; 388 AA.
XX
XX AAW73514;
XX 02-MAR-1999 (first entry)
DT Human TGFbetaRII:Fc protein sequence.
DE
DE
XX Transforming growth factor-beta receptor; TGF-beta receptor; arthritis;
KW fusion protein; fibroproliferative disorder; diabetic nephropathy;
KW glomerulonephritis; proliferative vitreoretinopathy; myelofibrosis;
KW collagen vascular disorder; therapy; human.
XX Homo sapiens.
OS
XX WO9848024-A1.
PN
XX 29-OCT-1998.
PD
XX 16-APR-1998; 98WO-US07587.
PF
XX 18-APR-1997; 97US-0044641.
PR
XX

Best Local Similarity 90.78; Pred. No. 7.3e-84;		Matches 235; Conservative 4; Mismatches 3; Indels 17; Gaps 2;	
QY	97	SSDE-----DEYEMSPNITDTMKFFLYVDKTHCPAPPELLGGPSVFLFPPKPK 148	
Db	139	STDECDNDHIFSEEYTTSSDPL-----VDKTHCPAPPELLGGPSVFLFPPKPK 189	
QY	149	DTLWISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTV 208	
Db	190	DTLWISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTV 249	
QY	209	LHODWLNKEYCKVSNKALPAPIEKTISKAKGQPREPOQVYTLPPSRDELTKNQVSTLCL 268	
Db	250	LHODWLNKEYCKVSNKALPAPIEKTISKAKGQPREPOQVYTLPPSRDELTKNQVSTLCL 309	
QY	269	VKGFYPSDIAVWESNGQPENNYKTPPVLDSDGSFFLYSLKLTVDKSRWQOGNVFSCSV 328	
Db	310	VKGFYPSDIAVWESNGQPENNYKTPPVLDSDGSFFLYSLKLTVDKSRWQOGNVFSCSV 369	
QY	329	HEALHNHYTQKSLSLSPGK 347	
Db	370	HEALHNHYTQKSLSLSPGK 388	
RESULT 15			
ID	AA97172	standard; Protein; 497 AA.	
XX	XX	AA97172;	
XX	XX	04-DEC-2000 (first entry)	
DE	XX	Human FGF-R1 Extracellular domain-Ig Fc fusion protein 3.	
KW	XX	FGF-R; fibroblast growth factor receptor; extracellular domain; IgG1; immunoglobulin; G1; oligomerization domain; Fc region; fusion protein; inhibitor; dimer; antagonist; cytostatic; anti-diabetic; vulnerary; ophthalmological; anti-proliferative.	
OS	XX	Homo sapiens.	
FH	XX	Key	
FT	FT	Peptide	
FT	FT	1..21	
FT	FT	/label= FGF-R1_signal_peptide	
FT	FT	22..257	
FT	FT	/label= FGF-R1_extracellular_domain	
FT	FT	/note= "the Ig I segment and acid box are deleted"	
FT	FT	59..111	
FT	FT	/label= Ig_II_segment	
FT	FT	157..222	
FT	FT	/label= Ig_III_segment	
FT	FT	258..265	
FT	FT	/label= Linker	
FT	FT	266..497	
FT	FT	/label= Human_IgG1_Fc_region	
FT	FT	/note= "Contains hinge region and domains CH2 and CH3"	
XX	XX	WO200046380-A2.	
XX	XX	10-AUG-2000.	
XX	XX	07-FEB-2000; 2000WO-US03166.	
XX	XX	08-FEB-1999; 99US-0119002.	
XX	XX	(CHIR) CHIRON CORP.	
XX	XX	Kavanaugh WM, Ballinger M;	
XX	XX	WPI; #2000-514961/46.	
XX	XX	N-PSDB; AAA52129.	
XX	XX	New polypeptide comprising a fibroblast growth factor receptor	

PT extracellular domain fused to a heterologous oligomerization domain for treating FGF-, angiogenesis-, or FGF receptor-mediated disorders

XX Claim 14; Page 58-59; 70pp; English.

XX Novel fusion protein constructs comprise a fibroblast growth factor (FGF) receptor (FGF-R) extracellular domain (ECD) lacking the immunoglobulin (Ig) I segment fused to a heterologous oligomerization domain that comprises an immunoglobulin Fc region, hinge region, CH1, CH2, CH3 or CH4 region, or light chain of an immunoglobulin molecule, or a peptide with a leucine zipper motif. The Ig I segment is not necessary for binding of acidic FGF and basic FGF (bFGF). The Ig I deletion further increases the affinity for aFGF and heparin, protects the core of the molecule from proteolysis, and abrogates the heparin requirement for aFGF binding. The new fusion polypeptides are better FGF inhibitors than FGF-R monomer proteins. The FGF-R-Ig Fc fusion dimers are active as FGF antagonists at subnanomolar concentrations and were 20-fold more potent than the FGF-R monomer protein as competitors of bFGF binding to immobilized FGF-Rs. The fusion constructs are useful to treat FGF-, angiogenesis-, or FGF-R-mediated disorders, such as tumorigenesis (e.g. bladder, breast, lung, rectal, testis and cervical tumours), neovascularization (e.g. diabetic retinopathy, neovascular glaucoma, wound healing and corneal scarring) and hyper-proliferation of vascular smooth muscle cells (e.g. postangioplasty and postatherectomy restenosis).

XX Sequence 497 AA;

SQ

Query Match 67.3%; Score 1249.5; DB 21; Length 497;

Best Local Similarity 75.8%; Pred. No. 1.2e-83;

Matches 247; Conservative 12; Mismatches 32; Indels 35; Gaps 5;

QY 46 PSNVPLKEVL-----WKKQDKVAELENSEFFRAFSFK-----NRVYLDTVSGSLTIYNL 95

Db 183 PDNLPPVQILKTAGVNTDTKEMEVHLRNVSEFDAGEYTCLAGNSIGLSHISAWLTV--L 240

QY 96 TSSDEDEYEMESPNITDTMKFFLYV-----DKTHTCPAPPELLGGPSVF 141

Db 241 EALEERPAVMTSP-----LYLEGSGSPGLOEPKSCDKTHTCPAPPELLGGPSVF 291

QY 142 LFPPKPKDTLMSIRTPETVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYR 201

Db 292 LFPPKPKDTLMSIRTPETVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYR 351

QY 202 VVSVLTVLHODWLNKEYCKVSNKALPAPIEKTISKAKGQPREPOQVYTLPPSRDELTKN 261

Db 352 VVSVLTVLHODWLNKEYCKVSNKALPAPIEKTISKAKGQPREPOQVYTLPPSRDELTKN 411

QY 262 QVSLTCLVKGFPSPDIAVWESNGQPENNYKTPPVLDSDGSFFLYSLKLTVDKSRWQOGN 321

Db 412 QVSLTCLVKGFPSPDIAVWESNGQPENNYKTPPVLDSDGSFFLYSLKLTVDKSRWQOGN 471

QY 322 VFSCVMHEALHNHYTQKSLSLSPGK 347

Db 472 VFSCVMHEALHNHYTQKSLSLSPGK 497

Search completed: January 28, 2003, 08:42:02

Job time : 38.7806 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 08:40:59 ; Search time 113.923 Seconds
(without alignments)
61.462 Million cell updates/sec

Title: US-09-730-465-8

Perfect score: 1856

Sequence: 1 MVAGSDAGRALGVLSVCLL.....MHEALHNHYTKSLSPGK 347

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1856	100.0	347	9	US-10-091-236-17
2	1856	100.0	347	10	US-09-796-033-8
3	1856	100.0	347	10	US-09-730-465-8
4	1251.5	67.4	388	10	US-09-734-300-9
5	1250.5	67.4	388	10	US-09-734-300-8
6	1249	67.3	859	9	US-09-935-868-7
7	1249	67.3	1158	9	US-09-935-868-26
8	1249	67.3	1168	9	US-09-935-868-24
9	1247	67.2	451	10	US-09-875-338-17
10	1247	67.2	698	10	US-09-875-338-9
11	1242	66.9	388	10	US-09-784-623-16
12	1242	66.9	492	10	US-09-845-899A-3
13	1240	66.8	399	9	US-09-832-659-2
14	1239.5	66.8	482	10	US-09-824-286-2
15	1238.5	66.7	691	9	US-09-935-868-20
16	1238.5	66.7	694	9	US-09-935-868-22
17	1238.5	66.7	780	9	US-09-935-868-34
18	1238.5	66.7	780	9	US-09-935-868-38
19	1238.5	66.7	780	9	US-09-935-868-42

20	1238	66.7	394	10	US-09-854-864-31	Sequence 31, Appl
21	1237	66.6	229	9	US-10-215-297-2	Sequence 2, Appl
22	1237	66.6	229	9	US-10-215-298-2	Sequence 2, Appl
23	1237	66.6	380	10	US-09-948-018-36	Sequence 36, Appl
24	1237	66.6	380	10	US-09-948-018-39	Sequence 39, Appl
25	1237	66.6	418	9	US-09-832-659-42	Sequence 42, Appl
26	1237	66.6	423	9	US-09-832-659-44	Sequence 44, Appl
27	1237	66.6	594	10	US-09-815-108-22	Sequence 22, Appl
28	1236	66.6	592	9	US-09-935-868-8	Sequence 8, Appl
29	1235	66.5	281	10	US-09-854-864-10	Sequence 10, Appl
30	1235	66.5	480	10	US-09-875-338-5	Sequence 5, Appl
31	1234	66.5	228	9	US-09-840-277-2	Sequence 2, Appl
32	1234	66.5	228	10	US-09-847-712-2	Sequence 2, Appl
33	1234	66.5	282	9	US-09-840-277-109	Sequence 109, App
34	1233.5	66.5	694	9	US-09-935-868-18	Sequence 18, Appl
35	1233.5	66.5	793	9	US-09-935-868-32	Sequence 32, Appl
36	1233	66.4	232	10	US-09-996-357-10	Sequence 10, Appl
37	1233	66.4	235	10	US-09-784-623-6	Sequence 6, Appl
38	1233	66.4	247	10	US-09-996-357-13	Sequence 13, Appl
39	1233	66.4	267	10	US-09-996-357-12	Sequence 12, Appl
40	1233	66.4	283	10	US-09-854-864-9	Sequence 9, Appl
41	1233	66.4	330	9	US-10-047-542-20	Sequence 20, Appl
42	1233	66.4	360	10	US-09-949-713-11	Sequence 11, Appl
43	1233	66.4	376	10	US-09-949-713-22	Sequence 22, Appl
44	1233	66.4	389	10	US-09-784-623-14	Sequence 14, Appl
45	1233	66.4	397	10	US-09-854-864-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-10-091-236-17
; Sequence 17, Application US/10091236
; Patent No. US20020168360A1
; GENERAL INFORMATION:

- ; APPLICANT: DINGIVAN, CHRISTINE A.
- ; TITLE OF INVENTION: METHODS OF PREVENTING OR TREATING INFLAMMATORY OR AUTOIMMUNE DISORDERS BY ADMINISTERING INTEGRIN ALPHA-V-BETA-3 ANTAGONISTS
- ; TITLE OF INVENTION: COMBINATION WITH OTHER PROPHYLACTIC OR THERAPEUTIC AGENTS
- ; FILE REFERENCE: 10271-053-999
- ; CURRENT APPLICATION NUMBER: US/10/091,236
- ; CURRENT FILING DATE: 2002-03-04
- ; PRIOR APPLICATION NUMBER: US 60/273,098
- ; PRIOR FILING DATE: 2001-03-02
- ; PRIOR APPLICATION NUMBER: US 60/316,321
- ; PRIOR FILING DATE: 2001-08-31
- ; NUMBER OF SEQ ID NOS: 17
- ; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 17
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-091-236-17

Query Match	100.0%	Score 1856;	DB 9;	Length 347;
Best Local Similarity	100.0%;	Pred. No. 1.7e-123;		
Matches 347;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Oy	1	MVAGSDAGRALGVLSVCLLHCFGFI	SCFSQOIYGVVYGNVTFHVPSNVPKLVKKQK	60
Db	1	MVAGSDAGRALGVLSVCLLHCFGFI	SCFSQOIYGVVYGNVTFHVPSNVPKLVKKQK	60
Oy	61	DKVAELNESEFRAFSKFNRYLDTV	SGSLTIYNLTSDEDEYEMESPNITDTMKFFLYV	120
Db	61	DKVAELNESEFRAFSKFNRYLDTV	SGSLTIYNLTSDEDEYEMESPNITDTMKFFLYV	120
Oy	121	DKTHTCCPAPFELGGPSVFLFPK	PDMLISRTPEVTCVVVDVSHEDPEVFNNYVD	180
Db	121	DKTHTCCPAPFELGGPSVFLFPK	PDMLISRTPEVTCVVVDVSHEDPEVFNNYVD	180
Oy	181	GVEVHNATKPREQYNSTYRVVSL	TVLHQDWLNGKEYKCKVSNKALPAPIETISKAK	240
Db	181	GVEVHNATKPREQYNSTYRVVSL	TVLHQDWLNGKEYKCKVSNKALPAPIETISKAK	240

Db 181 GVEVHNKTKPREEQNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 240
QY 241 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 300
Db 241 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 300
QY 301 DGSFFLYSKLTVDKSRWQGNVFSCVSMHEALHNHYTQKSLSLSPGK 347
Db 301 DGSFFLYSKLTVDKSRWQGNVFSCVSMHEALHNHYTQKSLSLSPGK 347

RESULT 2

US-09-796-033-8
; Sequence 8, Application US/09796033
; Patent No. US20020009446A1
; GENERAL INFORMATION:
; APPLICANT: Magilavy, Daniel
; TITLE OF INVENTION: METHOD OF MODULATING MEMORY EFFECTOR
; TITLE OF INVENTION: T-CELLS AND COMPOSITIONS
; FILE REFERENCE: 10274-044001
; CURRENT APPLICATION NUMBER: US/09/796,033
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: PCT/US99/20026
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: US 60/098,456
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(28)
US-09-796-033-8

Query Match 100.0%; Score 1856; DB 10; Length 347;
Best Local Similarity 100.0%; Pred. No. 1.7e-123;
Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVAGSDAGRALGVLSVVCLLHCFGFISCFISQIYGVVYGVNTHFVPSNVPLKEVLKKQK 60
Db 1 MVAGSDAGRALGVLSVVCLLHCFGFISCFISQIYGVVYGVNTHFVPSNVPLKEVLKKQK 60
QY 61 DKVAELENSEFRASFKNRVYLDTVSGSLTIYNTLTSSDEDEYEMESPNITDTMKFFLYV 120
Db 61 DKVAELENSEFRASFKNRVYLDTVSGSLTIYNTLTSSDEDEYEMESPNITDTMKFFLYV 120
QY 121 DKHTCCPPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 180
Db 121 DKHTCCPPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 180
QY 181 GVEVHNKTKPREEQNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 240
Db 181 GVEVHNKTKPREEQNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 240
QY 241 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 300
Db 241 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 300
QY 301 DGSFFLYSKLTVDKSRWQGNVFSCVSMHEALHNHYTQKSLSLSPGK 347
Db 301 DGSFFLYSKLTVDKSRWQGNVFSCVSMHEALHNHYTQKSLSLSPGK 347

RESULT 3

US-09-730-465-8
; Sequence 8, Application US/09730465
; Patent No. US20020009449A1
; GENERAL INFORMATION:
; APPLICANT: Wallner, Barbara P.
; Cooper, Kevin D.

; TITLE OF INVENTION: Method of Prophylaxis or Treatment of Antigen
; Presenting Cell Driven Skin Conditions Using
; Inhibitors of the CD2/LFA-3 Interaction
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/730,465
FILING DATE: 05-Dec-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/08755
FILING DATE: 06-OCT-1992
APPLICATION NUMBER: US 07/862,022
FILING DATE: 12-APR-1992
APPLICATION NUMBER: US 07/770,969
FILING DATE: 07-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Louis (PLM)
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: BGP-111CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 347 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-730-465-8

Query Match 100.0%; Score 1856; DB 10; Length 347;
Best Local Similarity 100.0%; Pred. No. 1.7e-123;
Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVAGSDAGRALGVLSVVCLLHCFGFISCFISQIYGVVYGVNTHFVPSNVPLKEVLKKQK 60
Db 1 MVAGSDAGRALGVLSVVCLLHCFGFISCFISQIYGVVYGVNTHFVPSNVPLKEVLKKQK 60
QY 61 DKVAELENSEFRASFKNRVYLDTVSGSLTIYNTLTSSDEDEYEMESPNITDTMKFFLYV 120
Db 61 DKVAELENSEFRASFKNRVYLDTVSGSLTIYNTLTSSDEDEYEMESPNITDTMKFFLYV 120
QY 121 DKHTCCPPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 180
Db 121 DKHTCCPPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 180
QY 181 GVEVHNKTKPREEQNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 240
Db 181 GVEVHNKTKPREEQNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 240
QY 241 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 300
Db 241 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 300
QY 301 DGSFFLYSKLTVDKSRWQGNVFSCVSMHEALHNHYTQKSLSLSPGK 347
Db 301 DGSFFLYSKLTVDKSRWQGNVFSCVSMHEALHNHYTQKSLSLSPGK 347

RESULT 4
US-09-734-300-9


```

: CURRENT FILING DATE: 2002-04-11
: PRIOR APPLICATION NUMBER: PCT/US99/220045
: PRIOR FILING DATE: 1999-09-22
: NUMBER OF SEQ ID NOS: 52
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 26
: LENGTH: 1158
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-935-868-26

```

Query Match 67.3%; Score 1249; DB 9; Length 1158;
Best Local Similarity 79.3%; Pred. No. 2.7e-80;
Matches 242; Conservative 11; Mismatches 30; Indels 22

Qy	65	BLENSERAPSGF-----KNRVYLDTVSGSLTYVNLTSDD-----EDEVEMESP	108
		: : : : : : : :	
Db	854	DVQNGFRINITYFTTIIGNETAANNVDSHRYETFLSLSATSLYVMVMAAYTDGGKQCP	913
Qy	109	NIDTNRKFFLY-----VDKXTHCPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCV	162
Db	914	EFTFTTPKFAQGEIESGDKXTHCPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCV	973
Qy	163	VVDVSHEDPEVFNNYVDGVEVINAKTKPREQVNSTYRVSVLTVLHQDLWNGREYCKC	222
Db	974	VVDVSHEDPEVFNNYVDGVEVINAKTKPREQVNSTYRVSVLTVLHQDLWNGREYCKC	1033
Qy	223	VSNKALPAPIETKISKARGOPPEQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWE	282
Db	1034	VSNKALPAPIETKISKAGQPPPEQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWE	1093
Qy	283	SNGOPENNYKTTPPVLDSDGSFFLYSKYTLVDKSRWQGNVPSCSVMHEALHHNYTKLSLS	342
Db	1094	SNGOPENNYKTTPPVLDSDGSFFLYSKYTLVDKSRWQGNVPSCSVMHEALHHNYTKLSLS	1153
Qy	343	LSPGK 347	
Db	1154	LSPGK 1158	

```

RESULT 8
US-09-935-868-24
; Sequence 24, Application US/09935868
; Patent No. US20020164690A1
; GENERAL INFORMATION:
; APPLICANT: Regeneron Pharmaceuticals, Inc
; TITLE OF INVENTION: Receptor Based Antagonists, and Methods of Making and Using
; FILE REFERENCE: REG 203D
; CURRENT APPLICATION NUMBER: US/09/935,868
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: PCT/US99/22045
; PRIOR FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24
; LENGTH: 1168
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-935-868-24

```

Query Match 67.3%; Score 1249; DB 9; Length 1168;
Best Local Similarity 79.3%; Pred. No. 2.8e-80;
Matches 242; Conservative 11; Mismatches 30; Indels 22; Gaps 3;

[illegible]

Db	984	VVDVSHEDPEVKFNWYVDGVEVINAKTKPREQYNTYRVVSVLTVLHQDMLNGKEYCK	1043
Qy	223	VSNKALPAPIEKTISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGYGSPDSIAVE	282
Db	1044	VSNKALPAPIEKTISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGYGSPDSIAVE	1103
Qy	283	SNQGPENNYKTTPTPVLDSDGSFFLYSKGLVDKSRMQGNVVFCSVMHEALHNHYTQKSLS	342
Db	1104	SNQGPENNYKTTPTPVLDSDGSFFLYSKGLVDKSRMQGNVVFCSVMHEALHNHYTQKSLS	1163
Qy	343	LSPGK 347	
Db	1164	LSPGK 1168	

RESULT 9
US-09-875-338-17
; Sequence 17, Application US/09875338
; Patent No. US2002095024A1
; GENERAL INFORMATION:
; APPLICANT: MIKESELL, GLEN E.
; APPLICANT: CHANG, HAN
; APPLICANT: FINGER, JOSHUA N.
; APPLICANT: YANG, GUCHEN
; APPLICANT: LU, PIN
; APPLICANT: ZHOU, XIA-DI
; APPLICANT: PEACH, ROBERT
; TITLE OF INVENTION: B7-RELATED NUCLEIC ACIDS AND POLYPEPTIDES USEFUL FOR
; TITLE OF INVENTION: IMMUNOMODULATION
; FILE REFERENCE: 3053-4071US2
; CURRENT APPLICATION NUMBER: US/09/875.338
; CURRENT FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/272,107
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/209,811
; PRIOR FILING DATE: 2000-06-06
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: fusion construct
; US-09-875-338-17

Query Match	67.2%	Score 1247;	DB 10;	Length 451;
Best Local Similarity	70.1%;	Pred. No. 1.3e-80;		
Matches 253;	Conservative	8;	Mismatches 40;	Indels 60; Gaps 6;
QY	33	ITYGVV-----YGNVTFHW-----PSNPVLKEVLAKK-----	58	
Db	105	ITYGVAMDYKYLTLKVASYRKINTHILKAVPETDEVELTQCATGYPLAEVSNPNVSVNPANT	164	
QY	59	-----QDKVAELENSFRASFSEKKNRVYLDITVSGSLTI--YNLTSSDEDEYEME	106	
Db	165	SHSRTPGELGYGVTSVURLKPPGCRNPFSC-----VVMNTHVRELTLASIDQLQSGMEPRTEFE	220	
QY	107	SNINITDTMKFFLYVDKTHTCPCCPAPELGGPSVFLFPKPKDLMISRTPEVTCVVVDV	166	
		:	:	
Db	221	PKS-----CDKTHTCPCCPAPELGGPSVFLFPKPKDLMISRTPEVTCVVVDV	270	
QY	167	SHEDPEVKFNMYVDGYEVENANAKTKPREQGYNSTYRVVSVLTVLHODWLNGLNGEYKCKVSNK	226	
Db	271	SHEDPEVKFNMYVDGYEVENANAKTKPREQGYNSTYRVVSVLTVLHODWLNGLNGEYKCKVSNK	330	
QY	227	ALPAPIETKISKAKGQRPQVYTLPPSRDELTKNQVSVLTCLVKGYFSDIAVWESNGQ	286	
Db	331	ALPAPIETKISKAKGQRPQVYTLPPSRDELTKNQVSVLTCLVKGYFSDIAVWESNGQ	390	
QY	287	PENNYKTPPVLDSGDSFFLYSKLTVDKSRWOOQGNFSCSVNHEALHNHYHTOKSLISLSPG	346	

```

*
Db 391 PENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKSLSPG 450
QY 347 K 347
Db 451 K 451

RESULT 10
US-09-875-338-9
; Sequence 9, Application US/09875338
; Patent No. US20020095024A1
; GENERAL INFORMATION:
; APPLICANT: MIKESELL, GLEN E.
; APPLICANT: CHANG, HAN
; APPLICANT: FINGER, JOSHUA N.
; APPLICANT: YANG, GUCHEN
; APPLICANT: LU, PIN
; APPLICANT: ZHOU, XIA-DI
; APPLICANT: PEACH, ROBERT
; TITLE OF INVENTION: B7-RELATED NUCLEIC ACIDS AND POLYPEPTIDES USEFUL FOR
; TITLE OF INVENTION: IMMUNOMODULATION
; FILE REFERENCE: 3053-4071US2
; CURRENT APPLICATION NUMBER: US/09/875,338
; CURRENT FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/272,107
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/209,811
; PRIOR FILING DATE: 2000-06-06
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 698
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: fusion construct
US-09-875-338-9

Query Match 67.2%; Score 1247; DB 10; Length 698;
Best Local Similarity 87.4%; Pred. No. 2.2e-80;
Matches 236; Conservative 5; Mismatches 17; Indels 12; Gaps 1;

QY 78 KNRVYLDVSGSLTYINLTSSDEDEYEMESPNTIDTMKFFLYVDKTHTCPCPAPELLGG 137
Db 441 RNPVQQDAGSVYITGQPTFPPEPKS-----CDKTHTCPCPAPELLGG 488

QY 138 PSVFLEPPPKDTLMISRTPEVTCVVVDVSHEDPEVKFNNYVDGVEVHNAKTKPRREQYN 197
Db 489 PSVFLEPPPKDTLMISRTPEVTCVVVDVSHEDPEVKFNNYVDGVEVHNAKTKPRREQYN 548

QY 198 STYRVSVLTVLHODWLNGEKYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDE 257
Db 549 STYRVSVLTVLHODWLNGEKYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDE 608

QY 258 LTRKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRW 317
Db 609 LTRKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRW 668

QY 318 QQGNVFCVSMHEALHNHYTKSLSPG 347
Db 669 QQGNVFCVSMHEALHNHYTKSLSPG 698

RESULT 11
US-09-784-623-16
; Sequence 16, Application US/09784623
; Patent No. US2002009454A1
; GENERAL INFORMATION:
; APPLICANT: Boone, Thomas C.
; APPLICANT: Hershenson, Susan
; APPLICANT: Bevilacqua, Michael P.
```

```

; APPLICANT: Collins, David S.
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR TREATING INFLAMMATORY
; TITLE OF INVENTION: DISEASES
; FILE REFERENCE: A-365F
; CURRENT APPLICATION NUMBER: US/09/784,623
; CURRENT FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: 09/131,247
; PRIOR FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: PCT/US 97/02131
; PRIOR FILING DATE: 1997-02-10
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Human
US-09-784-623-16

Query Match 66.9%; Score 1242; DB 10; Length 388;
Best Local Similarity 65.0%; Pred. No. 2.5e-80;
Matches 249; Conservative 16; Mismatches 30; Indels 88; Gaps 5;

QY 30 SQQIYGVVYGNVTFHVPSPNPLKQKQDKVAELENSEFRFSEKKNRYVLDTVSGS 89
Db 29 NQLVAGVLOG-----PNVLEE-----KIDVVPLE-----PHALFLGIHGK 65

QY 90 LTIYNLTSSDEDEYEMESPNTIDTM-----KFFLYV----- 120
Db 66 MCLSCVKSGDETRLQLEAVNITDLSENRKQDKRFAFIRSDSGPTTTFESAACPGWFLCTA 125

QY 121 -----DKTHTCCPCPAPELLGGPSVFLFP 144
Db 126 MEADQPVSLTNMPDEGVMTKFYQDEAAAEKPKSSDKTHTCCPCPAPELLGGPSVFLFP 185

QY 145 PKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNNYVDGVEVHNAKTKPREOYNSTYRVVS 204
Db 186 PPKDITLMISRTPEVTCVVVDVSHEDPEVKFNNYVDGVEVHNAKTKPREQNSTYRVVS 245

QY 205 VLTVLHODWLNGEKYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNOVS 264
Db 246 VLTVLHODWLNGEKYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNOVS 305

QY 265 LTLCKVGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFS 324
Db 306 LTLCKVGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFS 365

QY 325 CSVMHEALHNHYTKSLSPG 347
Db 366 CSVMHEALHNHYTKSLSPG 388

RESULT 12
US-09-845-899A-3
; Sequence 3, Application US/09845899A
; Patent No. US20020147326A1
; GENERAL INFORMATION:
; APPLICANT: CHAIKIN, MARGERY ANN
; APPLICANT: LYN, SALLY DOREEN PATRICIA
; APPLICANT: SWEET, RAYMOND W.
; APPLICANT: TRUNEH, ALEMSEGED
; TITLE OF INVENTION: HEXAMERIC FUSION PROTEINS AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: P50496
; CURRENT APPLICATION NUMBER: US/09/845,899A
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 09/202,346
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/043,948
; PRIOR FILING DATE: 1997-02-19
; PRIOR APPLICATION NUMBER: US 60/038,915
; PRIOR FILING DATE: 1997-02-21
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 3.0
```

```

; SEQ ID NO 3
; LENGTH: 492
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-845-B99A-3

Query Match      66.8%; Score 1242; DB 10; Length 492;
Best Local Similarity 59.2%; Pred. No. 3.6e-80;
Matches 250; Conservative 14; Mismatches 36; Indels 122; Gaps

QY 48 NVPLKE-----VLWKKQKDKVAELENSEFRAPSFKNRVYLDTVSG-SLIYVNLTSDDSD 101
   ||::|| :|:|:| | : : : :|:| | : | : | : | : | |
Db 53 NVSVEELAQTRIYWQEKKAIVLTWMSGDMINIWPYKRNTRIFDITNNLSIVILALRPSDEG 112
   ||
QY 102 EYE-----MESP----- 108
   ||
Db 113 TYECVWLKYEADFAKREHLAEVTLVKRADEPTPTSISDFEIPTSNIRRIICSTGGFPPEH 172
   ||
QY 109 -----NITDTMKFELVV----- 120
   |:|
Db 173 LSWLENGEELNAINTVSODPETELYAVSSKLDFNTTNSHFMCGLIKYGHRLVNQTFFNN 232
QY 121 -----DKHTCCPCPAPELLGGSVFLPFPKPOTLMISRTPEVTVCVVVD 165
Db 233 TTQOEHPDPQESAKDTHTCPPCPAPELLGGSVFLPFPKPOTLMISRTPEVTVCVVVD 292
QY 166 VSHEDPEVKFNWYVDCGEVHNAAKTPREEQYNSTRYRSVLTVLHQDLWLNCKEYKCVSN 225
   |||||||
Db 293 VSHEDPEVKFNWYVDCGEVHNAAKTPREEQYNSTRYRSVLTVLHQDLWLNCKEYKCVSN 352
QY 226 KALPAPTEKTIKSAKGOPREPQVYTLPPSRDELTKNOVSLCLVKGYEPSDIAVEHSNG 285
   |||||||
Db 353 KALPAPTEKTIKSAKGOPREPQVYTLPPSRDELTKNOVSLCLVKGYEPSDIAVEHSNG 412
QY 286 QPENNYKTPPVLDSDGSFFLYSKLTVDKRWQOQGVNFSCSVMHHEALHNHYTKQSLSLP 345
   |||||||
Db 413 QPENNYKTPPVLDSDGSFFLYSKLTVDKRWQOQGVNFSCSVMHHEALHNHYTKQSLSLA 472
QY 346 GK 347
   ||
Db 473 GK 474

RESULT 13
US-09-832-659-2
; Sequence 2, Application US/09832659
; Patent No. US20020155547A1
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; TITLE OF INVENTION: Interferon-Beta Fusion Proteins and Uses
; FILE REFERENCE: A064PCTSEQ
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US/09/832,659
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/120,237
; PRIOR FILING DATE: 1999-02-16
; PRIOR APPLICATION NUMBER: 60/104,491
; PRIOR FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 399
; TYPE: PRT
; ORGANISM: murine
US-09-832-659-2

Query Match      66.8%; Score 1240; DB 9; Length 399;
Best Local Similarity 86.3%; Pred. No. 3.6e-80;
Matches 233; Conservative 9; Mismatches 6; Indels 22; Gaps

QY 78 KNRVYLDTVSGSLTYINLTSSDEDEYEWESPNITDTMKFFLYVDKTHTCPPCPAPELLGG 137
   :| :: |:| | : : : :|:| | : | : | : | : | |
Db 152 RNRYFINLRATG-----YLNRNDOK-----VDKTHTCPPCPAPELLGG 189

; SEQ ID NO 3
; LENGTH: 492
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-824-286-2

Query Match      66.8%; Score 1239.5; DB 10; Length 482;
Best Local Similarity 100.0%; Pred. No. 4.8e-80;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps

QY 120 VKDTHTCPPCPAPELLGGSVFLPFPKPOTLMISRTPEVTVCVVDSHEDPEVKFNWY 179
Db 255 VKDTHTCPPCPAPELLGGSVFLPFPKPOTLMISRTPEVTVCVVDSHEDPEVKFNWY 314
```

```
Qy 180 DGEVHNAKTKPREQYNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKA 239
Db 315 DGEVHNAKTKPREQYNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKA 374
Qy 240 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTTPPVLD 299
Db 375 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTTPPVLD 434
Qy 300 SDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 347
Db 435 SDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 482

RESULT 15
US-09-935-868-20
; Sequence 20, Application US/09935868
; Patent No. US20020164690A1
; GENERAL INFORMATION:
; APPLICANT: Regeneron Pharmaceuticals, Inc
; TITLE OF INVENTION: Receptor Based Antagonists, and Methods of Making and Using
; FILE REFERENCE: REG 203D
; CURRENT APPLICATION NUMBER: US/09/935,868
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: PCT/US99/22045
; PRIOR FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-935-868-20

Query Match 66.7%; Score 1238.5; DB 9; Length 691;
Best Local Similarity 71.5%; Pred. No. 8.4e-80;
Matches 248; Conservative 10; Mismatches 38; Indels 51; Gaps 6;

Qy 39 GNVTFHV-----PSNVPLKEV-----LWKKQKD-----KVAELENSEFR 72
Db 358 GNLTVHTNVSDFLLLTWSNPYPDPDNYLNHLTYAVNINSENDPADFRINYVTVLEPSLRI 417
Qy 73 AFSSSEKNRY-YLDTVSGSLTIYNTLTSSD-----EDEYEMESPNITDTMKFELYV 120
Db 418 AASTLKGISYRARRAQAQSYNTWSENWSPSTKWHNSYREFFEQSG----- 464
Qy 121 DKHTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNMYVD 180
Db 465 DKHTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNMYVD 524
Qy 181 GVEVHNAKTKPREQYNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKA 240
Db 525 GVEVHNAKTKPREQYNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKA 584
Qy 241 GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTTPPVLD 300
Db 585 GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTTPPVLD 644
Qy 301 DGSSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 347
Db 645 DGSSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 691

Search completed: January 28, 2003, 08:53:01
Job time : 114.923 secs
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 08:38:45 ; Search time 15.6935 Seconds
(without alignments)
650.573 Million cell updates/sec

Title: us-09-730-465-8

Perfect score: 1856

Sequence: 1 MVAGSDAGRALGVLSVVCLL.....MHEALHNHYTKSLSLSPK 347

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents.AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1856	100.0	347	1 US-07-940-861-43	Sequence 43, Appl
2	1856	100.0	347	1 US-08-459-512-43	Sequence 43, Appl
3	1856	100.0	347	2 US-08-459-657-43	Sequence 43, Appl
4	1856	100.0	347	2 US-08-460-132-43	Sequence 43, Appl
5	1856	100.0	347	4 US-08-466-465-8	Sequence 8, Appl
6	1856	100.0	347	5 PCT-US92-02050-43	Sequence 43, Appl
7	1242	66.9	388	4 US-09-131-247-16	Sequence 16, Appl
8	1239.5	66.8	482	4 US-09-189-129-2	Sequence 2, Appl
9	1239	66.8	442	5 PCT-US96-10043-9	Sequence 9, Appl
10	1239	66.8	704	4 US-09-590-656-2	Sequence 9, Appl
11	1238.5	66.7	424	5 PCT-US95-03866-12	Sequence 12, Appl
12	1238.5	66.7	424	5 PCT-US95-03866-14	Sequence 14, Appl
13	1238.5	66.7	437	5 PCT-US96-10043-11	Sequence 11, Appl
14	1238	66.7	680	4 US-08-227-996C-15	Sequence 15, Appl
15	1235.5	66.6	488	4 US-08-776-511-2	Sequence 2, Appl
16	1234	66.5	552	1 US-08-243-010-6	Sequence 6, Appl
17	1233	66.4	232	2 US-08-595-043A-50	Sequence 50, Appl
18	1233	66.4	235	4 US-09-131-247-6	Sequence 6, Appl
19	1233	66.4	331	4 US-09-178-869-2	Sequence 2, Appl
20	1233	66.4	360	4 US-09-180-100-11	Sequence 11, Appl
21	1233	66.4	371	1 US-08-236-311-7	Sequence 7, Appl
22	1233	66.4	371	3 US-08-457-918-7	Sequence 7, Appl
23	1233	66.4	376	4 US-09-180-100-22	Sequence 22, Appl
24	1233	66.4	387	1 US-08-470-299-4	Sequence 4, Appl
25	1233	66.4	389	4 US-09-131-247-14	Sequence 14, Appl
26	1233	66.4	396	2 US-08-784-512-3	Sequence 3, Appl
27	1233	66.4	396	4 US-09-176-228-3	Sequence 3, Appl

Sequence 8, Appl
Sequence 7, Appl
Sequence 13, Appl
Sequence 7, Appl
Sequence 10, Appl
Sequence 4, Appl
Sequence 12, Appl
Sequence 8, Appl
Sequence 16, Appl
Sequence 16, Appl
Sequence 16, Appl
Sequence 16, Appl
Sequence 16, Appl
Sequence 18, Appl
Sequence 33, Appl
Sequence 22, Appl
Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-07-940-861-43
; Sequence 43, Application US/07940861
; Patent No. 5547853
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: WALLNER, Barbara P.
; APPLICANT: MILLER, Glenn T.
; APPLICANT: ROSA, Margaret D.
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 875 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022-6250
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/940,861
; FILING DATE: 21-OCT-1992
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02050
; FILING DATE: 12-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/667,971
; FILING DATE: 12-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,967
; FILING DATE: 07-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY, James F., Jr.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B151CIP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)715-0600
; TELEFAX: (212)715-0673
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```
; MOLECULE TYPE: protein
; us-07-940-861-43

Query Match      100.0%; Score 1856; DB 1; Length 347;
Best Local Similarity 100.0%; Pred. No. 3.2e-163;
Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVAGSDAGRALGVLSVVCLLHCFGIFSCFSQOIYGVVGNVTFHVPNSVPLKEVLKKQK 60
    |||
Db 1 MVAGSDAGRALGVLSVVCLLHCFGIFSCFSQOIYGVVGNVTFHVPNSVPLKEVLKKQK 60

Qy 61 DKVAELENSEFRAPSFKNRYVLDTVSGSLTIYNLTSSDEDEYEMESPNITDTMKFFLYV 120
    |||
Db 61 DKVAELENSEFRAPSFKNRYVLDTVSGSLTIYNLTSSDEDEYEMESPNITDTMKFFLYV 120

Qy 121 DKHTCPCPCAPPELLGPGSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 180
    |||
Db 121 DKHTCPCPCAPPELLGPGSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 180

Qy 181 GVEVHNKATKPREEQYNSTYRVVSVLTVLHQDLNKGKEYCKVSNKALPAPIETISKAK 240
    |||
Db 181 GVEVHNKATKPREEQYNSTYRVVSVLTVLHQDLNKGKEYCKVSNKALPAPIETISKAK 240

Qy 241 GQPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 300
    |||
Db 241 GQPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 300

Qy 301 DGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNNHYTQKLSLSPGK 347
    |||
Db 301 DGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNNHYTQKLSLSPGK 347

RESULT 2
us-08-459-512-43
; Sequence 43, Application US/08459512
; Patent No. 5728677
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: WALLNER, Barbara P.
; APPLICANT: MILLER, Glenn T.
; APPLICANT: ROSA, Margaret D.
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022-6250
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,512
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02050
; FILING DATE: 12-MAR-1992
; APPLICATION NUMBER: US 07/667,971
; FILING DATE: 12-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,967
; FILING DATE: 07-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY, James F., Jr.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B151CIP2
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: (212)715-0600
; TELEFAX: (212)715-0673
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-459-512-43

Query Match      100.0%; Score 1856; DB 1; Length 347;
Best Local Similarity 100.0%; Pred. No. 3.2e-163;
Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVAGSDAGRALGVLSVVCLLHCFGIFSCFSQOIYGVVGNVTFHVPNSVPLKEVLKKQK 60
    |||
Db 1 MVAGSDAGRALGVLSVVCLLHCFGIFSCFSQOIYGVVGNVTFHVPNSVPLKEVLKKQK 60

Qy 61 DKVAELENSEFRAPSFKNRYVLDTVSGSLTIYNLTSSDEDEYEMESPNITDTMKFFLYV 120
    |||
Db 61 DKVAELENSEFRAPSFKNRYVLDTVSGSLTIYNLTSSDEDEYEMESPNITDTMKFFLYV 120

Qy 121 DKHTCPCPCAPPELLGPGSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 180
    |||
Db 121 DKHTCPCPCAPPELLGPGSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 180

Qy 181 GVEVHNKATKPREEQYNSTYRVVSVLTVLHQDLNKGKEYCKVSNKALPAPIETISKAK 240
    |||
Db 181 GVEVHNKATKPREEQYNSTYRVVSVLTVLHQDLNKGKEYCKVSNKALPAPIETISKAK 240

Qy 241 GQPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 300
    |||
Db 241 GQPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 300

Qy 301 DGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNNHYTQKLSLSPGK 347
    |||
Db 301 DGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNNHYTQKLSLSPGK 347

RESULT 3
US-08-459-657-43
; Sequence 43, Application US/08459657
; Patent No. 5914111
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: WALLNER, Barbara P.
; APPLICANT: MILLER, Glenn T.
; APPLICANT: ROSA, Margaret D.
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 875 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022-6250
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,657
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02050
; FILING DATE: 12-MAR-1992
; APPLICATION NUMBER: US 07/667,971
; FILING DATE: 12-MAR-1991
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,967
; FILING DATE: 07-OCT-1991
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY, James F., Jr.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B151CIP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)715-0600
; TELEFAX: (212)715-0673
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-459-657-43

Query Match      100.0%; Score 1856; DB 2; Length 347;
Best Local Similarity 100.0%; Pred. No. 3.2e-163;
Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVAGSDAGRALGVLVVCLLHCFISCSQIYGVVYGNVTFHVPNSVPLKEVLKKQK 60
Db 1 MVAGSDAGRALGVLVVCLLHCFISCSQIYGVVYGNVTFHVPNSVPLKEVLKKQK 60
QY 61 DKVAELENSEFRASFKNRVYLDTVSGSLTIYNTLSSDEDEYEMESPNITDTMKFFLYV 120
Db 61 DKVAELENSEFRASFKNRVYLDTVSGSLTIYNTLSSDEDEYEMESPNITDTMKFFLYV 120
QY 121 DKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 180
Db 121 DKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 180
QY 181 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 240
Db 181 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 240
QY 241 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 300
Db 241 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 300
QY 301 DGSFFLYSKLTVDKSRWQGNVFSCVSMHEALHNHYTQKSLSLSPGK 347
Db 301 DGSFFLYSKLTVDKSRWQGNVFSCVSMHEALHNHYTQKSLSLSPGK 347

RESULT 4
US-08-460-132-43
; Sequence 43, Application US/08460132
; Patent No. 5928643
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: WALLNER, Barbara P.
; APPLICANT: MILLER, Glenn T.
; APPLICANT: ROSA, Margaret D.
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
; TITLE OF INVENTION: FUNCTION ASSOCIATED ANTIGEN 3
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 875 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022-6250
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/460,132
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/940,861
; FILING DATE: 21-OCT-1992
; APPLICATION NUMBER: PCT/US92/02050
; FILING DATE: 12-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/667,971
; FILING DATE: 12-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,967
; FILING DATE: 07-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY, James F., Jr.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B151CIP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)715-0600
; TELEFAX: (212)715-0673
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-460-132-43

Query Match      100.0%; Score 1856; DB 2; Length 347;
Best Local Similarity 100.0%; Pred. No. 3.2e-163;
Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVAGSDAGRALGVLVVCLLHCFISCSQIYGVVYGNVTFHVPNSVPLKEVLKKQK 60
Db 1 MVAGSDAGRALGVLVVCLLHCFISCSQIYGVVYGNVTFHVPNSVPLKEVLKKQK 60
QY 61 DKVAELENSEFRASFKNRVYLDTVSGSLTIYNTLSSDEDEYEMESPNITDTMKFFLYV 120
Db 61 DKVAELENSEFRASFKNRVYLDTVSGSLTIYNTLSSDEDEYEMESPNITDTMKFFLYV 120
QY 121 DKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 180
Db 121 DKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 180
QY 181 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 240
Db 181 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 240
QY 241 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 300
Db 241 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 300
QY 301 DGSFFLYSKLTVDKSRWQGNVFSCVSMHEALHNHYTQKSLSLSPGK 347
Db 301 DGSFFLYSKLTVDKSRWQGNVFSCVSMHEALHNHYTQKSLSLSPGK 347

RESULT 5
US-08-466-465-8
; Sequence 8, Application US/08466465
; Patent No. 6162432
; GENERAL INFORMATION:
; APPLICANT: Wallner, Barbara P.
; APPLICANT: Cooper, Kevin D.
; TITLE OF INVENTION: Method of Prophylaxis or Treatment of Antigen
; TITLE OF INVENTION: Presenting Cell Driven Skin Conditions Using
; TITLE OF INVENTION: Inhibitors of the CD2/LFA-3 Interaction
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
```

;; CITY: Boston
;; STATE: Massachusetts
;; COUNTRY: USA
;; ZIP: 02109-1875
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/466,465
;; FILING DATE:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US92/08755
;; FILING DATE: 06-OCT-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/862,022
;; FILING DATE: 12-APR-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/770,969
;; FILING DATE: 07-OCT-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Myers, Louis (PLM)
;; REGISTRATION NUMBER: 35,965
;; REFERENCE/DOCKET NUMBER: BGP-111CP
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (617)227-7400
;; TELEFAX: (617)227-5941
;; INFORMATION FOR SEQ ID NO: 8:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 347 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-466-465-8

Query Match 100.0%; Score 1856; DB 4; Length 347;
Best Local Similarity 100.0%; Pred. No. 3.2e-163;
Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVAGSDAGRALGVLSVWCLLHCFGFISCFSQIYGVVYGNVTFHVPSNVPLKEVLWKQK 60
DB 1 MVAGSDAGRALGVLSVWCLLHCFGFISCFSQIYGVVYGNVTFHVPSNVPLKEVLWKQK 60

QY 61 DKVAELENSEFRASFNNRVYLDTVSGSLTIYNTLTSDEDEYEMESPNITDTMKFFLYV 120
DB 61 DKVAELENSEFRASFNNRVYLDTVSGSLTIYNTLTSDEDEYEMESPNITDTMKFFLYV 120

QY 121 DKHTHTCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVYVDVSHEDPEVKFNMYVD 180
DB 121 DKHTHTCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVYVDVSHEDPEVKFNMYVD 180

QY 181 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDLNGLKEYCKVSNKALPAPIETISKAK 240
DB 181 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDLNGLKEYCKVSNKALPAPIETISKAK 240

QY 241 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPPVLD 300
DB 241 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPPVLD 300

QY 301 DGSFFLYSKLTVDKSRWQGNVFSCSVMHVHNAHNYHTQKLSLSPGK 347
DB 301 DGSFFLYSKLTVDKSRWQGNVFSCSVMHVHNAHNYHTQKLSLSPGK 347

RESULT 6

PCT-US92-02050-43
; Sequence 43, Application PC/TUS9202050
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: WALLNER, Barbara P.
; APPLICANT: MILLER, Glenn T.
; APPLICANT: ROSA, Margaret D.

;; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
;; TITLE OF INVENTION: FUNCTION ASSOCIATED ANTIGEN 3
;; NUMBER OF SEQUENCES: 43
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Fish & Neave
;; STREET: 875 Third Avenue
;; CITY: New York
;; STATE: New York
;; COUNTRY: U.S.A.
;; ZIP: 10022-6250
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US92/02050
;; FILING DATE: 19920312
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/667,971
;; FILING DATE: 12-MAR-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/770,967
;; FILING DATE: 07-OCT-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: HALEY, James F., Jr.
;; REGISTRATION NUMBER: 27,794
;; REFERENCE/DOCKET NUMBER: B151CIP2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212)715-0600
;; TELEFAX: (212)715-0673
;; TELEX: 14-8367
;; INFORMATION FOR SEQ ID NO: 43:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 347 amino acids
;; TYPE: AMINO ACID
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; PCT-US92-02050-43

Query Match 100.0%; Score 1856; DB 5; Length 347;
Best Local Similarity 100.0%; Pred. No. 3.2e-163;
Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVAGSDAGRALGVLSVWCLLHCFGFISCFSQIYGVVYGNVTFHVPSNVPLKEVLWKQK 60
DB 1 MVAGSDAGRALGVLSVWCLLHCFGFISCFSQIYGVVYGNVTFHVPSNVPLKEVLWKQK 60

QY 61 DKVAELENSEFRASFNNRVYLDTVSGSLTIYNTLTSDEDEYEMESPNITDTMKFFLYV 120
DB 61 DKVAELENSEFRASFNNRVYLDTVSGSLTIYNTLTSDEDEYEMESPNITDTMKFFLYV 120

QY 121 DKHTHTCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVYVDVSHEDPEVKFNMYVD 180
DB 121 DKHTHTCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVYVDVSHEDPEVKFNMYVD 180

QY 181 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDLNGLKEYCKVSNKALPAPIETISKAK 240
DB 181 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDLNGLKEYCKVSNKALPAPIETISKAK 240

QY 241 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPPVLD 300
DB 241 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPPVLD 300

QY 301 DGSFFLYSKLTVDKSRWQGNVFSCSVMHVHNAHNYHTQKLSLSPGK 347
DB 301 DGSFFLYSKLTVDKSRWQGNVFSCSVMHVHNAHNYHTQKLSLSPGK 347

RESULT 7

US-09-131-247-16
; Sequence 16, Application US/09131247


```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10043
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,213
; FILING DATE: 14-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lech, Karen F.
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 00786/284001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 442 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US96-10043-9

Query Match 66.8%; Score 1239; DB 5; Length 442;
Best Local Similarity 82.1%; Pred. No. 3.7e-106;
Matches 239; Conservative 8; Mismatches 12; Indels 32; Gaps 3;

Qy 77 FKNRYVLDVSGSLT-----IYNLT-----SSDEDEYEMESPNITDTMKF 116
Db 164 FPEPVTVNSNGALTPSGVHTFPVAVLQSSGLYSLSVTVTPSSDKVPEKS----- 214
Qy 117 FLYVDKTHPCPCPAPELLGSPSVFLFPKPKDTLWISRTPEVTCVVVDVSHEDPEVKFN 176
Db 215 ---CDKTHCPCPAPELLGSPSVFLFPKPKDTLWISRTPEVTCVVVDVSHEDPEVKFN 271
Qy 177 WYVDGVEVHNATKPREEQNSTYRYVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKT 236
Db 272 WYVDGVEVHNATKPREEQNSTYRYVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKT 331
Qy 237 SKAKGQPRQPVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP 296
Db 332 SKAKGQPRQPVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP 391
Qy 297 VLDSGSPFLSKLTVDKSRWQGNVFCVSMHEALHNHYTOKLSLSLSPGK 347
Db 392 VLDSGSPFLSKLTVDKSRWQGNVFCVSMHEALHNHYTOKLSLSLSPGK 442

RESULT 10
US-09-590-656-2
; Sequence 2, Application US/09590656
; Patent No. 6413932
; GENERAL INFORMATION:
; APPLICANT: Cerretti, Douglas P.
; APPLICANT: Borges, Luis G.
; APPLICANT: Fanslow, III, William C.
; TITLE OF INVENTION: TEK ANTAGONISTS
; FILE REFERENCE: 2900-A
; CURRENT APPLICATION NUMBER: US/09/590,656
; CURRENT FILING DATE: 2000-06-07
; PRIOR APPLICATION NUMBER: 60/137,889
; PRIOR FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 704
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-590-656-2

Query Match 66.8%; Score 1239; DB 4; Length 704;

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/03866
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,379
; FILING DATE: 28-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr, James F
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: CytoMed/2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-03866-12

Query Match 66.7%; Score 1238.5; DB 5; Length 424;
Best Local Similarity 82.5%; Pred. No. 3.9e-106;
Matches 236; Conservative 7; Mismatches 26; Indels 17; Gaps 2;

Qy 72 RAFSSFKNRVLDTVSGSLTIYNTLSDEDEYEMESPNITDTMKFFL-----YVD 121

```

Db 146 RSIDAFKDFVASETSDCVSVSTLSPKDSRVSVTKP-----FMLPPVAADPEPKSCD 198
QY 122 KTHCTPCPCAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDG 181
Db 199 KTHCTPCPCAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDG 258
QY 182 VEVHNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKG 241
Db 259 VEVHNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKG 318
QY 242 QPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSD 301
Db 319 QPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSD 378
QY 302 GSFFLYSKLTVDKSRWQOGNWFSCSVMHREALHNNHYTKQKLSLSPGK 347
Db 379 GSFFLYSKLTVDKSRWQOGNWFSCSVMHREALHNNHYTKQKLSLSPGK 424

RESULT 12

PCT-US95-03866-14
; Sequence 14, Application PC/TUS9503866
; GENERAL INFORMATION:
; APPLICANT: CytoMed, Inc. (all states except US)
; APPLICANT: Nocka, Karl (US only)
; APPLICANT: Lobell, Robert B (US only)
; TITLE OF INVENTION: STABILIZED DIMER OF KIT LIGAND AND
; TITLE OF INVENTION: FLT-3/FLK-2 LIGAND
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/03866
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,379
; FILING DATE: 28-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr, James F
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: CytoMed/2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-03866-14

Query Match 66.7%; Score 1238.5; DB 5; Length 424;
Best Local Similarity 82.5%; Pred. No. 3.9e-106;
Matches 236; Conservative 7; Mismatches 26; Indels 17; Gaps 2;
QY 72 RAFSSFNRYVLTVDVSGSLTIYNTSSDEDEYEMESNITDTMKFEL-----YVD 121
Db 146 RSIDAFKDFVASETSDCVSVSTLSPKDSRVSVTKP-----FMLPPVAADPEPKSCD 198
QY 122 KTHCTPCPCAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDG 181

Db 199 KTHCTPCPCAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDG 258
QY 182 VEVHNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKG 241
Db 259 VEVHNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKG 318
QY 242 QPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSD 301
Db 319 QPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSD 378
QY 302 GSFFLYSKLTVDKSRWQOGNWFSCSVMHREALHNNHYTKQKLSLSPGK 347
Db 379 GSFFLYSKLTVDKSRWQOGNWFSCSVMHREALHNNHYTKQKLSLSPGK 424

RESULT 13

PCT-US96-10043-11
; Sequence 11, Application PC/TUS9610043
; GENERAL INFORMATION:
; APPLICANT: the General Hospital Corporation
; TITLE OF INVENTION: P-SELECTIN LIGANDS AND RELATED MOLECULES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02210-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10043
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,213
; FILING DATE: 14-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lech, Karen F.
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 00786/284001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US96-10043-11

Query Match 66.7%; Score 1238.5; DB 5; Length 437;
Best Local Similarity 78.8%; Pred. No. 4.1e-106;
Matches 238; Conservative 10; Mismatches 19; Indels 35; Gaps 3;
QY 49 VPLKEVYLV---NRKQDKQKVAELEENSEFRAPSSFNRYVLTVDVSGSLTIYNTSSDEDEYEM 105
Db 168 IPKSDVWVYTDWKKDCKEPLKQKEKER-----KQEGESDP 203
QY 106 ESPNITDTMKFELYVDKTHCTPCPCAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVD 165
Db 204 EGEPKS-----CDKTHCTPCPCAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVD 255

```
QY 166 VSHEDPEVKFNMYVDCGEVHNKTKPREQYNSTYRVSVLTVLHODWLNKKEYKCKVSN 225
|||||
Db 256 VSHDEPEVKFNMYVDCGEVHNKTKPREQYNSTYRVSVLTVLHODWLNKKEYKCKVSN 315
|||||
QY 226 KALPAPIEKTISKAKQPREQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVEMESNG 285
|||||
Db 316 KALPAPIEKTISKAKQPREQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVEMESNG 375
|||||
QY 286 QPENNNKTYTPPVLDSDGSFFLYSKLTVDKSRMOQGNVFCVSMHEALHNHYTOKLSLSP 345
|||||
Db 376 QPENNNKTYTPPVLDSDGSFFLYSKLTVDKSRMOQGNVFCVSMHEALHNHYTOKLSLSP 435
|||||
QY 346 GK 347
||
Db 436 GK 437

RESULT 14
US-08-227-496C-15
; Sequence 15, Application US/08227496C
; Patent No. 6130202
; GENERAL INFORMATION:
; APPLICANT: Greve, Jeffrey M.
; APPLICANT: McClelland, Alan
; TITLE OF INVENTION: Multimeric Forms of Human
; TITLE OF INVENTION: Rhinovirus Receptor Protein
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bayer Corporation
; STREET: 400 Morgan Lane
; CITY: West Haven
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06516
; COMPUTER READABLE FORM:
; MEDIUM TYPE: diskette, 1.44 Mb storage
; COMPUTER: Dell Optiplex GX1
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect 8.0 for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/227,496C
; FILING DATE: 04/14/94
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/903,069
; FILING DATE: 06/22/92
; APPLICATION NUMBER: 07/704,984
; FILING DATE: 05/24/91
; APPLICATION NUMBER: 07/556,238
; FILING DATE: 07/20/90
; ATTORNEY/AGENT INFORMATION:
; NAME: Barbara A. Shlmei
; REGISTRATION NUMBER: 29,862
; REFERENCE/DOCKET NUMBER: MTI 214.2C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203) 812-2786
; TELEFAX: (203) 812-5492
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 680 amino acid residues
; TYPE: amino acids
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: protein
; HYPOTHETICAL: no
; FRAGMENT TYPE: complete sequence
; FEATURE:
; NAME/KEY: tICAM(185)/IgG fusion protein
; OTHER INFORMATION: amino acid residues 1-453 =
; OTHER INFORMATION: tICAM(453); amino acid residues 454-680 = amino
; OTHER INFORMATION: acid residues 216-442 of human IgG1 heavy chain
US-08-227-496C-15
```

```
Query Match 66.7%; Score 1238; DB 4; Length 680;
Best Local Similarity 99.6%; Pred. No. 8.6e-106;
Matches 228; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 119 YVDKTHHTCCPPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNMY 178
|||||
Db 452 YEDKTHHTCCPPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNMY 511
|||||
QY 179 VDCGEVHNKTKPREQYNSTYRVSVLTVLHODWLNKKEYKCKVSNKALPAPIEKTISK 238
|||||
Db 512 VDCGEVHNKTKPREQYNSTYRVSVLTVLHODWLNKKEYKCKVSNKALPAPIEKTISK 571
|||||
QY 239 AKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVEMESNGQPENNYKTTTPPV 298
|||||
Db 572 AKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVEMESNGQPENNYKTTTPPV 631
|||||
QY 299 DSDGSFFLYSKLTVDKSRMOQGNVFCVSMHEALHNHYTOKLSLSPGK 347
|||||
Db 632 DSDGSFFLYSKLTVDKSRMOQGNVFCVSMHEALHNHYTOKLSLSPGK 680
|||||

RESULT 15
US-08-776-511-2
; Sequence 2, Application US/08776511
; Patent No. 6153190
; GENERAL INFORMATION:
; APPLICANT: Young, Peter R.
; APPLICANT: Erickson-Miller, Connie
; TITLE OF INVENTION: Method for Obtaining Receptor Agonist
; TITLE OF INVENTION: Antibodies
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation- Corporate
; ADDRESSEE: Patents
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,511
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Jervis, Herbert H.
; REGISTRATION NUMBER: 31,171
; REFERENCE/DOCKET NUMBER: SBC P50349-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5015
; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 488 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-776-511-2

Query Match 66.6%; Score 1235.5; DB 4; Length 488;
Best Local Similarity 92.5%; Pred. No. 9e-106;
Matches 234; Conservative 1; Mismatches 11; Indels 7; Gaps 2;

QY 95 LTSDEDEYEWESPNITDTMKFFLYVDKTHTCPPCAPPELLGGPSVFLFPPPKDPTLMIS 154
|||||
Db 243 LTPSDLOPIEGRG---TEPKS----ADKTHTCPPCAPPELLGGPSVFLFPPPKDPTLMIS 295
|||||
QY 155 RTPPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNKTKPREQYNSTYRVSVLTVLHODW 214
|||||
```

Db 296 RPEVTCVVVDYSHEDPEVKENWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWL 355
Qy 215 NGKEYCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYP 274
Db 356 NGKEYCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYP 415
Qy 275 SDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSLKLTVDKSRWQOGNMFSCSVMEALHN 334
Db 416 SDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSLKLTVDKSRWQOGNMFSCSVMEALHN 475
Qy 335 HYTKQSLSLSPCK 347
Db 476 HYTKQSLSLSPCK 488

Search completed: January 28, 2003, 08:40:11
Job time : 17.6935 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 08:49:45 ; Search time 14 Seconds
(without alignments)
109,868 Million cell updates/sec

Title: US-09-730-465-2_COPY_50_65

Perfect score: 84

Sequence: 1 PLKEVLWKQKDKVAE 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 2770

Minimum DB seq length: 0

Maximum DB seq length: 16

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	33.3	12	2 S69123	proton-translocati
2	27	32.1	14	2 I54945	gene C protein - E
3	25	29.8	8	2 A39308	glycine reductase
4	24	28.6	13	2 P00445	urotensin II - lau
5	24	28.6	15	2 S36888	ribosomal protein
6	24	28.6	15	2 E49037	Tcr delta chain V-
7	23	27.4	15	2 S21240	alpha-glucosidase
8	23	27.4	15	2 S21202	glucan 1,4-alpha-g
9	23	27.4	15	2 B56891	gamma 2 gliadin -
10	22	26.2	16	2 S03405	hydrogenase (EC 1.
11	21	25.0	9	2 A60356	118K stomach cance
12	21	25.0	13	2 S60046	early nodulin 40 -
13	21	25.0	14	1 QMWAPP	polistes mastopara
14	21	25.0	15	2 S33781	acetolactate synth
15	21	25.0	16	2 G45681	orf 61.1 - phage T
16	20	23.8	10	2 E49033	T-cell receptor ga
17	20	23.8	10	2 PA0116	ferredoxin-NADP re
18	20	23.8	11	2 PQ0731	unidentified 5.7/3
19	20	23.8	12	2 A26093	microbial collagen
20	20	23.8	12	2 J50424	urotensin II-B pep
21	20	23.8	14	2 C60414	somatostatin - sil
22	20	23.8	14	2 B60842	somatostatin I - c
23	20	23.8	14	2 A60622	somatostatin - spo
24	20	23.8	14	2 A60840	somatostatin I - E
25	20	23.8	14	2 S00172	somatostatin I - s
26	20	23.8	14	2 B83836	hypothetical prote
27	20	23.8	16	2 A28144	ribosomal protein
28	20	23.8	16	2 H29501	fibrinopeptide A -
29	19	22.6	12	2 I58273	thyroglobulin - ra

30 19 22.6 13 2 B35245 histone H1.c - mou
31 19 22.6 13 2 A35245 histone H1a - mous
32 19 22.6 14 2 A61308 hemocyanin chain 2
33 19 22.6 15 2 PNO118 hemoglobin beta ch
34 19 22.6 15 2 PA0102 fructose-bisphosph
35 19 22.6 15 2 PH1788 T cell receptor al
36 19 22.6 15 2 A45103 7 alpha-hydroxy-4-
37 19 22.6 16 2 E37290 homeotic protein G
38 18.5 22.0 14 2 S14336 mastoparan B - hor
39 18 21.4 10 2 F49033 T-cell receptor ga
40 18 21.4 11 1 LFTWME probable trpEG lea
41 18 21.4 12 2 I64829 gene HEXA protein
42 18 21.4 13 2 A32734 enkephalin precurs
43 18 21.4 13 2 E39778 lactase phosphotra
44 18 21.4 14 2 JH0328 probursin tetradece
45 18 21.4 14 2 PH1626 Ig H chain V-D-J r

ALIGNMENTS

RESULT 1

S69123

proton-translocating transhydrogenase - Rhodospirillum rubrum (fragment)

C:Species: Rhodospirillum rubrum

C:Date: 10-Mar-1998 #sequence_revision 24-Apr-1998 #text_change 24-Apr-1998

C:Accession: S69123

R:Diggle, C.; Hutton, M.; Jones, G.R.; Thomas, C.M.; Jackson, J.B.

Eur. J. Biochem. 228, 719-726, 1995

A:Title: Properties of the soluble polypeptide of the proton-translocating transhydro

A:Reference number: S69123; MUID:95255277; PMID:7737169

A:Accession: S69123

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-12 <DIG>

Query Match 33.3%; Score 28; DB 2; Length 12;

Best Local Similarity 50.0%; Pred. No. 2.5e+02;

Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 EVLWKQK 11

:||| |:

Db 3 DVVWKVQR 10

RESULT 2

I54945

gene C protein - Escherichia coli (fragment)

C:Species: Escherichia coli

C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 08-Oct-1999

C:Accession: I54945

R:Tao, T.; Bourne, J.C.; Blumenthal, R.M.

J. Bacteriol. 173, 1367-1375, 1991

A:Title: A family of regulatory genes associated with type II restriction-modificatio

A:Reference number: I54945; MUID:91139577; PMID:1995598

A:Accession: I54945

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-14 <RES>

A:Cross-references: GB:M63619; NID:g147664; PIDN:AAA24555.1; PID:g147665

Query Match 32.1%; Score 27; DB 2; Length 14;

Best Local Similarity 45.5%; Pred. No. 4.3e+02;

Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 LKEVLWKQK 12

||||:|:

Db 1 LKEVIMKKHE 11

RESULT 3

A93308

glycine reductase (EC 1.4.99.-) sulfhydryl protein C, alpha chain - Clostridium stick

C:Species: Clostridium sticklandii
C:Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 15-Aug-1997
C:Accession: A39308
R:Stadtman, T.C.; Davis, J.N.
J. Biol. Chem. 266, 22147-22153, 1991
A:Title: Glycine reductase protein C. Properties and characterization of its role in the
A:Reference number: A39308; MUID:92042141; PMID:1939235
A:Accession: A39308
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-8 <SFA>
C:Function:
C:Description: glycine reductase complex catalyzes the reductive deamination of glycine
C:Keywords: ATP; oxidoreductase

Query Match 29.8%; Score 25; DB 2; Length 8;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LKEVLW 7
:| |||
Db 1 MKPVLW 6

RESULT 4
PQ0445
urotensin II - laughing frog
C:Species: Rana ridibunda (laughing frog)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 03-Mar-1995
C:Accession: PQ0445
R:Conlon, J.M.; O'Harte, F.; Smith, D.D.; Tonon, M.C.; Vaudry, H.
Biochem. Biophys. Res. Commun. 188, 578-583, 1992
A:Title: Isolation and primary structure of urotensin II from the brain of a tetrapod, t
A:Reference number: PQ0445; MUID:93075134; PMID:1445302
A:Accession: PQ0445
A:Molecule type: protein
A:Residues: 1-13 <CON>
A:Experimental source: brain
C:Superfamily: urotensin II

Query Match 28.6%; Score 24; DB 2; Length 13;
Best Local Similarity 57.1%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 LKEVLWK 8
:| |||
Db 4 LSECFWK 10

RESULT 5
S36888
ribosomal protein S12 - Mycobacterium bovis (fragment)
C:Species: Mycobacterium bovis
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 31-Oct-1997
C:Accession: S36888
R:Ohara, N.; Kimura, M.; Higashi, Y.; Yamada, T.
FEBS Lett. 331, 9-14, 1993
A:Title: Isolation and amino acid sequence of the 30S ribosomal protein S19 from Mycobac
A:Reference number: S36887; MUID:94009653; PMID:8405418
A:Accession: S36888
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-15 <OH>
C:Superfamily: Escherichia coli ribosomal protein S12
C:Keywords: protein biosynthesis; ribosome

Query Match 28.6%; Score 24; DB 2; Length 15;
Best Local Similarity 23.1%; Pred. No. 1.4e+03;
Matches 3; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

QY 2 LKEVLWKQDKV 14
:| |||
Db 2 LKEVLWKQDKV 15

RESULT 6
E49037
TCR delta chain V-D-J region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C:Accession: E49037
R:Ezquerria, A.; Wilde, D.B.; McConnell, T.J.; Sturmhofel, K.; Valas, R.B.; Shevach, E.
Eur. J. Immunol. 22, 491-498, 1992
A:Title: Mouse autoreactive gamma/delta T cells. II. Molecular characterization of th
A:Reference number: A49037; MUID:92164730; PMID:1311262
A:Accession: E49037
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-15 <EQ>
A:Cross-references: GB:S90653; NID:g246296; PIDN:AAB21551.1; PID:g246297
A:Experimental source: dendritic epidermal T-cell lines
A>Note: sequence extracted from NCBI backbone (NCBIN:90653, NCBIP:90665)

Query Match 28.6%; Score 24; DB 2; Length 15;
Best Local Similarity 27.3%; Pred. No. 1.4e+03;
Matches 3; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 6 LWKKOKDKVAE 16
:| ||| :|
Db 3 LWEYRRDTTDD 13

RESULT 7
S21240
alpha-glucosidase (EC 3.2.1.20) I - Bacillus "thermoamyloliquefaciens" (fragment)
C:Species: Bacillus "thermoamyloliquefaciens"
C:Date: 22-Nov-1993 #sequence_revision 01-Nov-1996 #text_change 21-Aug-1998
C:Accession: S21240
R:Suzuki, Y.; Yonezawa, K.; Hattori, M.; Takii, Y.
Eur. J. Biochem. 205, 249-256, 1992
A:Title: Assignment of Bacillus thermoamyloliquefaciens KP1071 alpha-glucosidase I to
nce and in structural parameters calculated from the amino acid composition.
A:Reference number: S21202; MUID:92209510; PMID:1555585
A:Accession: S21240
A:Molecule type: protein
A:Residues: 1-15 <SUZ>
A:Experimental source: strain KP1071
C:Superfamily: alpha-glucosidase; alpha-amylase core homology
C:Keywords: glycosidase; hydrolase

Query Match 27.4%; Score 23; DB 2; Length 15;
Best Local Similarity 37.5%; Pred. No. 2e+03;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 LKEVLWKK 9
:| |||
Db 1 MKKAWKE 8

RESULT 8
S21202
glucan 1,4-alpha-glucosidase (EC 3.2.1.3) - Bacillus stearothermophilus (fragment)
N:Alternate names: exo-alpha-1-4-glucosidase I
C:Species: Bacillus stearothermophilus
C:Date: 22-Nov-1993 #sequence_revision 01-Nov-1996 #text_change 21-Aug-1998
C:Accession: S21202
R:Suzuki, Y.; Yonezawa, K.; Hattori, M.; Takii, Y.
Eur. J. Biochem. 205, 249-256, 1992
A:Title: Assignment of Bacillus thermoamyloliquefaciens KP1071 alpha-glucosidase I to
nce and in structural parameters calculated from the amino acid composition.
A:Reference number: S21202; MUID:92209510; PMID:1555585
A:Accession: S21202
A:Molecule type: protein
A:Residues: 1-15 <SUZ>
A:Experimental source: ATCC 12016
C:Superfamily: alpha-glucosidase; alpha-amylase core homology

C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 27.4%; Score 23; DB 2; Length 15;
Best Local Similarity 37.5%; Pred. No. 2e+03;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 LKEVLWKK 9
-|: ||:
Db 1 MKKTWKE 8

RESULT 9

B56891

gamma 2 gliadin - wheat (fragment)

C;Species: Triticum sp. (wheat)

C;Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 08-Dec-1995

C;Accession: B56891

R;Sjostrom, H.; Friis, S.U.; Noren, O.; Anthonsen, D.

Clin. Chim. Acta 207, 227-237, 1992

A;Title: Purification and characterisation of antigenic gliadins in coeliac disease.

A;Reference number: A56891; MUID:93009000; PMID:1395028

A;Contents: Kadett

A;Accession: B56891

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-15 <S70>

A;Note: sequence extracted from NCBI backbone (NCBIP:119388)

Query Match 27.4%; Score 23; DB 2; Length 15;
Best Local Similarity 40.0%; Pred. No. 2e+03;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 PLKEVLWKK 10
|: ||: |:
Db 6 PSGOVQWPQ 15

RESULT 10

S03405

hydrogenase (EC 1.18.99.1) small chain - Alcaligenes eutrophus (fragment)

C;Species: Alcaligenes eutrophus

C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 13-Sep-1998

C;Accession: S03405

R;Lorenz, B.; Schneider, K.; Kratzin, H.; Schlegel, H.G.

Biochim. Biophys. Acta 995, 1-9, 1989

A;Title: Immunological comparison of subunits isolated from various hydrogenases of aerob

A;Reference number: S03404; MUID:89166625; PMID:2493816

A;Accession: S03405

A;Molecule type: protein

A;Residues: 1-16 <LOR>

A;Experimental source: strain H16, DMS 541

C;Superfamily: hydrogenase (Nife) small chain

C;Keywords: hydrogen metabolism; iron-sulfur protein; membrane bound; metalloprotein; ni

Query Match 26.2%; Score 22; DB 2; Length 16;
Best Local Similarity 57.1%; Pred. No. 3.1e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 PLKEVLW 7
|: ||:
Db 5 PRTPVLW 11

RESULT 11

A60356

H18K stomach cancer antigen - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 17-Mar-1999

C;Accession: A60356

R;Shiraishi, Y.

Int. J. Cancer 45, 783-787, 1990

A;Title: Western blotting analysis for malignant lymphoma and stomach cancer antigens fr

A;Reference number: A60356; MUID:90216080; PMID:2323853

A;Accession: A60356
A;Molecule type: protein
A;Residues: 1-9 <SHI>
C;Keywords: glycoprotein

Query Match 25.0%; Score 21; DB 2; Length 9;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLKEVL 6
||: |:
Db 2 PLKPV 7

RESULT 12

S60046

early nodulin 40 - spring vetch

N;Alternate names: hypothetical protein 2

C;Species: Vicia sativa (spring vetch, tare)

C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 21-Jul-2000

C;Accession: S60046; TI0953

R;Viñ, I.; Yang, W.C.; Pallisgaard, N.; Ostergaard Jensen, E.; van Kammen, A.; Bissel

Plant Mol. Biol. 28, 1111-1119, 1995

A;Title: VSENOD5, VSENOD12 and VSENOD40 expression during Rhizobium-induced nodule fo

A;Reference number: S60044; MUID:96011756; PMID:7548828

A;Accession: S60046

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-13 <VIJ>

A;Cross-references: EMBL:X83683; NID:gl167893; PIDN:CAB37926.1; PID:g4468051

R;Christiansen, A.; Hansen, A.C.; Viñ, I.; Pallisgaard, N.; Larsen, K.; Yang, W.C.; B

submitted to the EMBL Data Library, December 1995

A;Description: A novel type of DNA binding protein interacts with a conserved sequenc

A;Reference number: Z17228

A;Accession: TI0953

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-4 <CHR>

A;Cross-references: EMBL:X95995; NID:gl360633; PID:e225825

Query Match 25.0%; Score 21; DB 2; Length 13;
Best Local Similarity 37.5%; Pred. No. 3.7e+03;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 LKEVLWKK 9
|: ||:
Db 1 MKLICWQK 8

RESULT 13

QMWAPP

polistes mastoparan - paper wasp (Polistes jadwigae)

C;Species: Polistes jadwigae

C;Date: 24-Sep-1981 #sequence_revision 24-Sep-1981 #text_change 08-Dec-1995

C;Accession: A01780

R;Hirai, Y.; Ueno, Y.; Yasuhara, T.; Yoshida, H.; Nakajima, T.

Biomed. Res. 1, 185-187, 1980

A;Title: A new mast cell degranulating peptide, polistes mastoparan, in the venom of

A;Reference number: A01780

A;Accession: A01780

A;Molecule type: protein

A;Residues: 1-14 <HIR>

C;Comment: This cytotoxic peptide from wasp venom induces mast cell degranulation.

C;Superfamily: mastoparan

C;Keywords: amidated carboxyl end

F;14/Modified site: amidated carboxyl end (Leu) #status experimental

Query Match 25.0%; Score 21; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 3.9e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WKK 9
|||

Db 3 WKK 5

RESULT 14

S33781
acetylactate synthase (EC 4.1.3.18) large chain, valine-sensitive - Serratia marcescens
C:Species: Serratia marcescens
C>Date: 19-Mar-1997 #sequence_revision 24-Mar-1999 #text_change 24-Mar-1999
C:Accession: S33781
R:Yang, J.H.; Kim, S.S.
Biochim. Biophys. Acta 1157, 178-184, 1993
A:Title: Purification and characterization of the valine sensitive acetylactate synthase
A:Reference number: S33781; MUID:93283409; PMID:8507653
A:Accession: S33781
A:Molecule type: protein
A:Residues: 1-15 <YAN>
A:Experimental source: ATCC 25419
C:Complex: heterotetramer; two small and two large chains
C:Function:
A:Description: catalyzes the condensation of pyruvate and alpha-ketobutyrate to form alpha-ketoglutarate
A:Pathway: valine, leucine, and isoleucine biosynthesis
A>Note: this isoenzyme exhibits homotropic allosterism with pyruvate
C:Keywords: branched-chain amino acid biosynthesis; carbon-carbon lyase; flavoprotein; H

Query Match 25.0%; Score 21; DB 2; Length 15;
Best Local Similarity 50.0%; Pred. No. 4.2e+03;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 7 WKKQKD 12
| : |
Db 9 WARQLD 14

RESULT 15

G45681
orf 61.1 - phage T6 (fragment)
C:Species: phage T6
C>Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994
C:Accession: G45681
R:Selick, H.E.; Stormo, G.D.; Dyson, R.L.; Alberts, B.M.
J. Virol. 67, 2305-2316, 1993
A:Title: Analysis of five presumptive protein-coding sequences clustered between the pri
A:Reference number: A45681; MUID:93188183; PMID:8383243
A:Accession: G45681
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-16 <SEL>
A>Note: sequence extracted from NCBI backbone (NCBIP:128349)

Query Match 25.0%; Score 21; DB 2; Length 16;
Best Local Similarity 66.7%; Pred. No. 4.5e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 8 KKQKDK 13
| : |
Db 4 KKKKSK 9

Search completed: January 28, 2003, 08:54:57
Job time : 16 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 08:44:25 ; Search time 10 seconds
(without alignments)
66.362 Million cell updates/sec

Title: US-09-730-465-2_COPY_50_65

Perfect score: 84

Sequence: 1 PLKEVLWKQKDKVAE 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 822

Minimum DB seq length: 0

Maximum DB seq length: 16

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	23	27.4	15	1 LCK_DROME	P81829 drosophila
2	23	27.4	15	1 MALU_BACTQ	P80072 bacillus th
3	21	25.0	13	1 NO40_PEA	P55959 pisum sativ
4	21	25.0	13	1 NO40_VICSA	P55961 vicia sativ
5	21	25.0	13	1 RPOC_MYCGA	P47716 mycoplasma
6	21	25.0	14	1 MAST_PARID	P42716 parapolybia
7	21	25.0	14	1 MAST_POLJA	P01517 polistes ja
8	20	23.8	12	1 UR2B_CATCO	P04559 catostomus
9	20	23.8	12	1 UR2B_CYPCA	P04561 cyprinus ca
10	20	23.8	12	1 UR2_POLSP	P81022 polyodon sp
11	20	23.8	14	1 SMS_MYOSC	P20750 myoxocephal
12	20	23.8	14	1 SMS_ALLMI	P31885 alligator m
13	18.5	22.0	14	1 MAST_VESBA	P21654 vespa basal
14	18	21.4	10	1 SPI_HALRO	Q10997 halocynthia
15	18	21.4	11	1 LPW_THETH	P05624 thermus the
16	18	21.4	12	1 NO40_SESRO	O24369 sesbania ro
17	18	21.4	15	1 PGKH_PHPYA	P80659 physcomitre
18	18	21.4	16	1 FIBA_MELME	P14456 melles melles
19	18	21.4	16	1 FIBA_MUSVI	P14458 musceta vis
20	17	20.2	7	1 WMA2_ACHFU	P35920 achatina fu
21	17	20.2	7	1 WMA3_ACHFU	P35921 achatina fu
22	17	20.2	9	1 LPCA_STAAU	P36884 staphylococ
23	17	20.2	11	1 CEPI_ACHFU	P22790 achatina fu
24	17	20.2	11	1 O20A_COMTE	P80464 comamonas t
25	17	20.2	12	1 NO40_LOTJA	O22426 lotus japon
26	17	20.2	12	1 RS19_CLYEP	Q46490 clover yell
27	17	20.2	12	1 RS19_TOBBP	Q36251 tomato big
28	17	20.2	12	1 UR2A_CATCO	P04558 catostomus
29	17	20.2	12	1 UR2_GILMI	P01147 gillichthys
30	17	20.2	12	1 UR2_SCYCA	P35490 scylliorhnu
31	17	20.2	13	1 AU11_LITRA	P82386 littoria ran
32	17	20.2	13	1 AU12_LITRA	P82387 littoria ran
33	17	20.2	14	1 RS19_PRUPA	Q44160 prunus arme

34	17	20.2	15	1 CLOA_RAT	P31720 rattus norv
35	17	20.2	15	1 CYSK_CLOPA	P81340 clostridium
36	17	20.2	15	1 THL_CLOPA	P81347 clostridium
37	17	20.2	15	1 UC27_MAIZE	P80633 zea mays (m
38	17	20.2	16	1 MMPX_SOLTU	P80501 solanum tub
39	16	19.0	9	1 COW_CONVE	P83047 conus ventr
40	16	19.0	9	1 NEF_HV128	P12481 human immun
41	16	19.0	9	1 ULAD_HUMAN	P31929 homo sapien
42	16	19.0	10	1 APE_CAPGI	P80474 capnocytoph
43	16	19.0	10	1 GON1_PETMA	P04378 petromyzon
44	16	19.0	10	1 GON3_PETMA	P30948 petromyzon
45	16	19.0	12	1 CALM_TETTH	Q05055 tetrahymena

ALIGNMENTS

RESULT 1

ID	LCK_DROME	STANDARD;	PRT;	15 AA.
AC	P81829;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Leucokinin (DLK).			
GN	PP OR DLK.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;			
OC	Insecta; pterygota; Neoptera; Endopterygota; Diptera; Brachycera;			
OC	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_taxID=7227;			
[1]				
RN	SEQUENCE.			
RP	TISSUE=Neurosecretory cell;			
RC	MEDLINE=20044845; PubMed=10574744;			
RX	Terhzaz S., O'Connell F.C., Pollock V.P., Kean L., Davies S.A.,			
RA	Veenstra J.A., Dow J.A.T.;			
RA	"Isolation and characterization of a leucokinin-like peptide of			
RT	Drosophila melanogaster."			
RL	J. Exp. Biol. 202:3667-3676(1999).			
[2]				
RN	SEQUENCE FROM N.A.			
RP	STRAIN=Berkley;			
RC	Celniker S.E., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G.,			
RA	Butenhoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,			
RA	Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,			
RA	Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim E., Lee B.,			
RA	Lewis S., Li P., Lomotan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,			
RA	Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,			
RA	Sethi H., Shrir E., Swirskas R.R., Wan K.H., Weinburg T., Zhang R.,			
RA	Zieran L.B., Rubin G.M.;			
RL	Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.			
CC	!- FUNCTION: ACTS THROUGH INTRACELLULAR CALCIUM IN MALPIGHIAN TUBULE			
CC	STELLATE CELLS TO RAISE CHLORIDE CONDUCTANCE.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL: AC006496; -; NOT ANNOTATED_CDS.			
DR	FlyBase: FBgn0028418; Leucokinin.			
KW	Neuropeptide; Amidation.			
FT	MOD_RES 15 15			
SQ	SEQUENCE 15 AA; 1743 MW; 4793A08F251C9525 CRC64;			

Query Match 27.4%; Score 23; DB 1; Length 15;
Best Local Similarity 71.4%; Pred. No. 7.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Query Match 25.0%; Score 21; DB 1; Length 13;
Best Local Similarity 37.5%; Pred. No. 1.3e+03;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 LKEVLWKK 9
: | : | : |
Db 1 MKLLCWQK 8

RESULT 5

RPOC_MYCGA
ID RPOC_MYCGA STANDARD; PRT; 13 AA.
AC P47716;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA-directed RNA polymerase beta' chain (EC 2.7.7.6) (Transcriptase
beta' chain) (RNA polymerase beta' subunit) (Fragment).
GN RPOC.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A5969Var.B;
RA Skamrov A.V., Rozovskaya T.A., Goldman M.A., Feoktistova E.S.,
RA Beabekashvili R.S.;
RL Submitted (XXX-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
SUBSTRATES.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
(RNA)(N).
CC -!- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
BETA' CHAIN.
CC -!- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC -----
DR EMBL; L38402; AAB40952.1; -.
KW Transferase; DNA-directed RNA polymerase; Transcription.
FT NON_TER 13
SQ SEQUENCE 13 AA; 1630 MW; 4BEC27C7480D4333 CRC64;

Query Match 25.0%; Score 21; DB 1; Length 13;
Best Local Similarity 66.7%; Pred. No. 1.3e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 8 KKQKDK 13
: | : | : |
Db 6 KKNKK 11

RESULT 6

MAST_PARID
ID MAST_PARID STANDARD; PRT; 14 AA.
AC P42716;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Mastoparan.
OS Parapolybia indica.
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
OC Aculeata; Vespoidea; Vespidae; Polistinae; Parapolybia.

OX NCBI_TaxID=31921;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RA Toki T., Yasuhara T., Nakajima T.;
RT "Isolation and sequential analysis of peptides on the venom sac of
RT Parapolybia indica.";
RL Eisei Dobutsu 39:105-111(1988).
CC -!- FUNCTION: MAST CELL DEGRANULATING PEPTIDE. ACTIVATES G PROTEINS
CC THAT COUPLE TO PHOSPHOLIPASE C.
KW Mast cell degranulation; Venom; Amidation.
FT MOD_RES 14 14
SQ SEQUENCE 14 AA; 1619 MW; CA376CD3BA6D80DD CRC64;

Query Match 25.0%; Score 21; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 WKK 9
: | : | : |
Db 3 WKK 5

RESULT 7

MAST_POLJA
ID MAST_POLJA STANDARD; PRT; 14 AA.
AC P01517;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Polistes mastoparan.
OS Polistes jadwigae (Paper wasp).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
OC Aculeata; Vespoidea; Vespidae; Polistinae; Polistes.
OX NCBI_TaxID=7457;
RN [1]
RP SEQUENCE.

RC TISSUE=Venom;
RA Hirai Y., Ueno Y., Yasuhara T., Yoshida H., Nakajima T.;
RT "A new mast cell degranulating peptide, polistes mastoparan, in the
RT venom of Polistes jadwigae.";
RL Biomed. Res. 1:185-187(1980).
CC -!- FUNCTION: MAST CELL DEGRANULATING PEPTIDE. ACTIVATES G PROTEINS
CC THAT COUPLE TO PHOSPHOLIPASE C.

DR PIR; A01780; QMWAPP.
KW Mast cell degranulation; Venom; Amidation.
FT MOD_RES 14 14
SQ SEQUENCE 14 AA; 1636 MW; 26472A5BF4778D8 CRC64;

Query Match 25.0%; Score 21; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 WKK 9
: | : | : |
Db 3 WKK 5

RESULT 8

UR2B_CATCO
ID UR2B_CATCO STANDARD; PRT; 12 AA.
AC P04559;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Urotensin IIB (U-IIB) (UIIB).
OS Catostomus commersoni (White sucker).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Catostomidae; Catostomus.
OX NCBI_TaxID=7971;
RN [1]


```

RX MEDLINE=89065329; PubMed=2904391;
RA Conlon J.M., Deacon C.F., Hazon N., Henderson I.W., Thim L.;
RT "Somatostatin-related and glucagon-related peptides with unusual
RL structural features from the European eel (Anguilla anguilla).";
RL Gen. Comp. Endocrinol. 72:181-189(1988).
CC -|- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
DR PIR: S00172; S00172.
DR PIR: B60842; B60842.
DR PIR: A60840; A60840.
DR InterPro: IPR004250; Somatostatin.
DR Pfam: PF03002; Somatostatin; 1.
KW Hormone; Multigene family.
FT DISULFID 3
FT SEQUENCE 14 AA; 1640 MW; D6270F5C09682679 CRC64;

Query Match 23.8%; Score 20; DB 1; Length 14;
Best Local Similarity 50.0%; Pred. No. 2.le+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 KEVLWK 8
DB 4 KNFEWK 9

RESULT 12
SMS_ALLMI STANDARD; PRT; 14 AA.
AC P31885;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Somatostatin-14
OS Alligator mississippiensis (American alligator), and
OS Trachemys scripta (red-eared slider turtle) (Pseudemys scripta);
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylidae; Alligatorinae; Alligator.
OX NCBI_TaxID=8496, 34903;
RN [1]
RP SEQUENCE.
RC SPECIES-T.scripta;
RX MEDLINE=90341082; PubMed=1974347;
RA Conlon J.M., Hicks J.W.;
RT "Isolation and structural characterization of insulin, glucagon and
RT somatostatin from the turtle, Pseudemys scripta.";
RL Peptides 11:461-466(1990).
CC -|- FUNCTION: SOMATOSTATIN INHIBITS THE RELEASE OF SOMATOTROPIN.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- SIMILARITY: BELONGS TO THE SOMATOSTATIN FAMILY.
DR PIR: C60414; C60414.
DR InterPro: IPR004250; Somatostatin.
DR Pfam: PF03002; Somatostatin; 1.
KW Hormone.
FT DISULFID 3
FT SEQUENCE 14 AA; 1640 MW; D6270F5C09682679 CRC64;

Query Match 23.8%; Score 20; DB 1; Length 14;
Best Local Similarity 50.0%; Pred. No. 2.le+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 KEVLWK 8
DB 4 KNFEWK 9

```

```

RESULT 13
MST_VESBA STANDARD; PRT; 14 AA.
ID MAST_VESBA
AC P21634;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Mastoparan B.
DE Vespa basalis (Hornet).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
OC Aculeata; Vespoidea; Vespidae; Vespinae; Vespa.
OX NCBI_TaxID=7444;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=91174755; PubMed=2006909;
RA Lo C.-L., Hwang L.-L.;
RT "Structure and biological activities of a new mastoparan isolated
RL from the venom of the hornet Vespa basalis.";
RL Biochem. J. 274:453-456(1991).
CC -|- FUNCTION: MAST CELL DEGRANULATING PEPTIDE. ACTIVATES G PROTEINS
CC THAT COUPLE TO PHOSPHOLIPASE C.
DR PIR: S14336; S14336.
KW Mast cell degranulation; Venom; Amidation.
FT MOD_RES 14
FT SEQUENCE 14 AA; 1613 MW; D35944CA193A19A2 CRC64;

Query Match 22.0%; Score 18.5; DB 1; Length 14;
Best Local Similarity 40.0%; Pred. No. 3.6e+03;
Matches 4; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 2 LKEVL-WKKQ 10
DB 3 LKSIVSWAKK 12

RESULT 14
SPI_HALRO STANDARD; PRT; 10 AA.
ID SPI_HALRO
AC Q10997;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Serine proteinase inhibitor (fragment).
OS Halocynthia roretzi (Sea squirt).
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Pyuridae; Halocynthia.
OX NCBI_TaxID=7729;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph; PubMed=8759295;
RX MEDLINE=96321313; PubMed=8759295;
RA Shishikura F., Abe T., Ohtake S.-I., Tanaka K.;
RT "Purification and characterization of a 58,000-Da proteinase
RT inhibitor from the hemolymph of a solitary ascidian, Halocynthia
RT roretzi.";
RL Comp. Biochem. Physiol. 114B:1-9(1996).
CC -|- FUNCTION: STRONGLY INHIBITS TRYPSIN AND PLASMA ENZYME(S) ACTIVITY.
CC -|- SUBUNIT: MONOMER.
CC -|- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR InterPro: IPR000215; Serpin.
DR PROSITE: PS00284; SERPIN; PARTIAL.
KW Serpin; Serine protease inhibitor; Glycoprotein; Plasma.
FT NON_TER 10
FT SEQUENCE 10 AA; 1104 MW; 4225C73B1B187AA3 CRC64;

Query Match 21.4%; Score 18; DB 1; Length 10;
Best Local Similarity 50.0%; Pred. No. 3e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 8 KQKDKVA 15

```

```
Db      | ::|||
        3 KGECKVA 10

RESULT 15
LPW_THETH
ID LPW_THETH STANDARD; PRT; 11 AA.
AC P05624;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Trp operon leader peptide.
GN TRPL.
OS Thermus thermophilus.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Thermales;
OC Thermaceae; Thermus.
OX NCBI_TaxID=274;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HB8 / ATCC 27634;
RX MEDLINE=89000781; PubMed=2844259;
RA Sato S., Nakada Y., Kanaya S., Tanaka T.:
RT "Molecular cloning and nucleotide sequence of Thermus thermophilus
RL HB8 trpE and trpG.";
RL Biochim. Biophys. Acta 950:303-312(1988).
CC -!- FUNCTION: THIS PROTEIN IS INVOLVED IN CONTROL OF THE BIOSYNTHESIS
CC OF TRYPTOPHAN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X07744; CAA30565.1; -.
DR PIR; S03315; LFTWWE.
KW Tryptophan biosynthesis; Leader peptide.
SQ SEQUENCE 11 AA; 1228 MW; 364B295A772DC5A7 CRC64;

Query Match 21.4%; Score 18; DB 1; Length 11;
Best Local Similarity 50.0%; Pred. No. 3.4e+03;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 LKEVLW 7
Db      | |||
        3 LPSALW 8

Search completed: January 28, 2003, 08:54:00
Job time : 12 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model
Run on: January 28, 2003, 08:49:05 ; Search time 29 Seconds
(without alignments)
113.681 Million cell updates/sec

Title: US-09-730-465-2_COPY_50_65
Perfect score: 84
Sequence: 1 PLKEVLMKKQDKVAE 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 3565

Minimum DB seq length: 0
Maximum DB seq length: 16

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phase:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_virus:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	34.5	15	5 P82207	P82207 bombyx mori
2	27	32.1	14	2 Q47599	Q47599 escherichia
3	26	31.0	15	4 Q9UCU7	Q9UCU7 homo sapien
4	25	29.8	15	3 Q9URE0	Q9URE0 saccharomyc
5	24.5	29.2	13	2 Q9R8R9	Q9R8R9 streptococc
6	24	28.6	15	2 Q9R347	Q9R347 mycobacteri
7	23	27.4	15	2 Q9R5L9	Q9R5L9 bacillus st
8	23	27.4	15	10 Q9S8V2	Q9S8V2 triticum ae
9	22	26.2	12	5 O61574	O61574 ostertagia
10	21	25.0	12	10 Q93WF2	Q93WF2 lupinus lut
11	21	25.0	15	2 Q05991	Q05991 staphylococ
12	21	25.0	15	2 Q9R586	Q9R586 serratia ma
13	21	25.0	15	6 Q9GLJ0	Q9GLJ0 sus scrofa
14	21	25.0	16	11 Q9CTB2	Q9CTB2 mus musculu
15	20	23.8	9	4 Q9UWA0	Q9UWA0 homo sapien
16	20	23.8	11	2 Q56413	Q56413 escherichia

17	20	23.8	11	4 Q9UE69	Q9UE69 homo sapien
18	20	23.8	11	4 Q9H4H5	Q9H4H5 homo sapien
19	20	23.8	13	5 Q9W5Q6	Q9W5Q6 drosophila
20	20	23.8	14	16 Q9KCS9	Q9KCS9 bacillus ha
21	20	23.8	15	1 Q9UWH4	Q9UWH4 pyrococcus
22	20	23.8	16	4 Q9BQF2	Q9BQF2 homo sapien
23	20	23.8	16	10 Q94F61	Q94F61 triticum ae
24	19	22.6	9	2 Q93E20	Q93E20 streptococc
25	19	22.6	10	13 Q9PRU1	Q9PRU1 cynops pyrr
26	19	22.6	12	2 Q8VLX8	Q8VLX8 thermus the
27	19	22.6	12	10 Q93WB7	Q93WB7 lupinus lut
28	19	22.6	12	11 Q63579	Q63579 rattus norv
29	19	22.6	13	12 Q83171	Q83171 cauliflowe
30	19	22.6	14	2 Q9R303	Q9R303 chlamydia t
31	19	22.6	15	2 Q9S0V1	Q9S0V1 nitrogen fi
32	19	22.6	15	4 Q9UBK0	Q9UBK0 homo sapien
33	19	22.6	15	15 Q8UM88	Q8UM88 human immun
34	19	22.6	16	4 Q9NZH9	Q9NZH9 homo sapien
35	19	22.6	16	4 Q9UD41	Q9UD41 homo sapien
36	19	22.6	16	11 Q9QW74	Q9QW74 mus sp. hom
37	18	21.4	9	2 P82568	P82568 streptococc
38	18	21.4	10	2 Q8RIT1	Q8RIT1 anaplasma p
39	18	21.4	10	4 Q8WTT4	Q8WTT4 homo sapien
40	18	21.4	10	8 Q9XMB4	Q9XMB4 aegilops ta
41	18	21.4	11	2 Q77569	Q77569 escherichia
42	18	21.4	11	7 Q77895	Q77895 oreochromis
43	18	21.4	12	2 Q9R5F7	Q9R5F7 helicobacte
44	18	21.4	12	4 Q9UMZ8	Q9UMZ8 homo sapien
45	18	21.4	12	13 Q9PVC7	Q9PVC7 esox lucius

ALIGNMENTS

RESULT 1

P82207 PRELIMINARY; PRT; 15 AA.
AC P82207
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DE Unknown protein from 2D-page (Fragment).
OS Bombyx mori (Silk moth)
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Bombycoidea; Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
RN [1]
RP SEQUENCE.
RC STRAIN=XINHANG X KEMING; TISSUE=BODY WALL, AND FAT BODY;
RX MEDLINE=21177481; PubMed=11280594;
RA Zhong B.X.;
RT "protein database for several tissues derived from five instar of silkworm.";
RL I Chuan Hsueh Pao 28:217-224(2001).
FT NON_TER 15
SQ SEQUENCE 15 AA; 1877 MW; 580FGBD4703CA70C CRC64;

Query Match 34.5%; Score 29; DB 5; Length 15;
Best Local Similarity 35.7%; Score No. 5,6e+02;
Matches 5; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

RESULT 2

Q47599 PRELIMINARY; PRT; 14 AA.
ID Q47599
AC Q47599;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

```

DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE C. (Fragment).
GN C.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OC NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91139577; PubMed=1995588;
RA Tao T., Bourne J.C., Blumenthal R.M.;
RT "A family of regulatory genes associated with type II restriction-
RT modification systems.";
RL J. Bacteriol. 173:1367-1375(1991).
DR EMBL; M63619; AAA24555.1; -.
FT NON_TER 1
SQ SEQUENCE 14 AA; 1705 MW; 77B6CA60581A4F3B CRC64;

Query Match 32.1%; Score 27; DB 2; Length 14;
Best Local Similarity 45.3%; Pred. No. 1.1e+03;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LKEVLWKQKD 12
Db 1 LKEVIMEKKHE 11

RESULT 3
Q9UCC7 ID Q9UCC7 PRELIMINARY; PRT; 15 AA.
AC Q9UCC7;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Midkine (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=94059921; PubMed=8241100;
RA Novotny W.F., Maffi T., Mehta R.L., Milner P.G.;
RT "Identification of novel heparin-releasable proteins, as well as the
RT cytokines midkine and pleiotrophin, in human postheparin plasma.";
RL Arterioscler. Thromb. 13:1798-1805(1993).
SQ SEQUENCE 15 AA; 1527 MW; C34B6B97878474AC CRC64;

Query Match 31.0%; Score 26; DB 4; Length 15;
Best Local Similarity 83.3%; Pred. No. 1.7e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 9 KQKDV 14
Db 3 KKKDV 8

RESULT 4
Q9URE0 ID Q9URE0 PRELIMINARY; PRT; 15 AA.
AC Q9URE0;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Type II topoisomerase, topoisomerase II (Fragment).
GN TOP2 AND YNL088W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95130576; PubMed=7829529;

```

```

RA Elsea S.H., Hsiung Y., Nitiss J.L., Osheroff N.;
RL J. Biol. Chem. 270:1913-1920(1995).
RG SGD; S0005032; TOP2.
FT NON_TER 1
FT NON_TER 15
SQ SEQUENCE 15 AA; 2094 MW; 0A6A37F6B81B85F6 CRC64;

Query Match 29.8%; Score 25; DB 3; Length 15;
Best Local Similarity 40.0%; Pred. No. 2.4e+03;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 7 WKXQDKVAE 16
Db 2 YQRRKDYMSE 11

RESULT 5
Q9R8R9 ID Q9R8R9 PRELIMINARY; PRT; 13 AA.
AC Q9R8R9;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE C5a peptidase (Fragment).
GN SCPA.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=AP1;
RX MEDLINE=98298075; PubMed=9632622;
RA Barge A., Rasmussen M., Bjorck L.;
RT "Identification of an insertion sequence located in a region encoding
RT virulence factors of Streptococcus pyogenes.";
RL Infect. Immun. 66:3449-3453(1998).
DR EMBL; AF064540; AAC38768.1; -.
FT MEROPS; S08.020; -.
FT NON_TER 13
SQ SEQUENCE 13 AA; 1603 MW; 5EDADCDA6CEB723 CRC64;

Query Match 29.2%; Score 24.5; DB 2; Length 13;
Best Local Similarity 63.6%; Pred. No. 2.5e+03;
Matches 7; Conservative 1; Mismatches 0; Indels 3; Gaps 1;

Qy 8 KKQK---DKVA 15
Db 3 KKQKLPFDKLA 13

RESULT 6
Q9R547 ID Q9R547 PRELIMINARY; PRT; 15 AA.
AC Q9R547;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE 30S ribosomal protein S12 homolog (Fragment).
OS Mycobacterium bovis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE.
RX MEDLINE=94009653; PubMed=8405418;
RA Ohara N., Kimura M., Higashi Y., Yamada T.;
RT "Isolation and amino acid sequence of the 30S ribosomal protein S19
RT from Mycobacterium bovis BCG.";
RL FEBS Lett. 331:9-14(1993).
SQ SEQUENCE 15 AA; 1808 MW; 788302A51B653234 CRC64;

Query Match 28.6%; Score 24; DB 2; Length 15;

```

```
Best Local Similarity 23.1%; Pred. No. 3.4e+03;
Matches 3; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

QY 2 LKEVLWKKQDKV 14
Db 3 IQQLVRKGRDRI 15

RESULT 7
Q9R5L9 PRELIMINARY; PRT; 15 AA.
AC Q9R5L9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE EXO-alpha-1,4-glucosidase (EC 3.2.1.20) (Fragment).
OS Bacillus stearothermophilus
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE.
RX MEDLINE=92209510; PubMed=1555585;
RA Suzuki Y., Yonezawa K., Hattori M., Takii Y.;
RT "Assignment of Bacillus thermoamyloliquefaciens KP1071 alpha-
glucosidase I to an exo-alpha-1,4-glucosidase, and its striking
similarity to bacillary oligo-1,6-glucosidases in N-terminal sequence
RT and in structural parameters calculated from the amino acid
composition."
RL Eur. J. Biochem. 205:249-256(1992).
SQ SEQUENCE 15 AA; 1931 MW; 6284CE40013D3042 CRC64;

Query Match 27.4%; Score 23; DB 2; Length 15;
Best Local Similarity 37.5%; Pred. No. 5e+03;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 LKEVLWKK 9
Db 1 MKKTWKE 8

RESULT 8
Q9S8V2 PRELIMINARY; PRT; 15 AA.
AC Q9S8V2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE Gamma 2 gliadin (Fragment).
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE.
RX MEDLINE=93009000; PubMed=1395028;
RA Sjostrom H., Friis S.U., Noren O., Anthonsen D.;
RT "Purification and characterisation of antigenic gliadins in coeliac
disease."
RL Clin. Chim. Acta 207:227-237(1992).
SQ SEQUENCE 15 AA; 1742 MW; 2B5C8365ACC8B32E CRC64;

Query Match 27.4%; Score 23; DB 10; Length 15;
Best Local Similarity 40.0%; Pred. No. 5e+03;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 PLKEVLWKKQ 10
Db 6 PSGQVQWQQ 15

RESULT 9
Q91574 PRELIMINARY; PRT; 12 AA.
AC Q91574;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 1.4 kDa protein (Fragment).
OS Ostertagia ostertagi.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Trichostrongyloidea; Haemonchidae; Ostertagiinae; Ostertagia.
OX NCBI_TaxID=6317;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20233682; PubMed=10769181;
RA Moore J., Tetley L., Devaney E.;
RT "Identification of abundant mRNAs from the third stage larvae of the
parasitic nematode, Ostertagia ostertagi."
RL Biochem. J. 347:763-770(2000).
DR EMBL; AF052049; AAC06297.1; -.
KW Hypothetical protein.
FT NON_TER 12
SQ SEQUENCE 12 AA; 1369 MW; C6FF867C1CA776C6 CRC64;

Query Match 26.2%; Score 22; DB 5; Length 12;
Best Local Similarity 55.6%; Pred. No. 5.7e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 8 KKQDKVAE 16
Db 4 KQPSKRAE 12

RESULT 10
Q93WF2 PRELIMINARY; PRT; 12 AA.
AC Q93WF2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Early nodulin.
GN ENOD40B.
OS Lupinus luteus (Yellow lupine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Genisteae; Lupinus.
OX NCBI_TaxID=3873;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. VENTUS;
RA Podkowinski J., Grabowska B., Kisiel A., Dlugaszewska B.,
RA Nimmagadda S.;
RT "ENOD40B from Lupinus luteus."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF352375; AAK51422.1; -.
DR EMBL; AF352372; AAK51419.1; -.
SQ SEQUENCE 12 AA; 1403 MW; 283958AE7CB326C3 CRC64;

Query Match 25.0%; Score 21; DB 10; Length 12;
Best Local Similarity 50.0%; Pred. No. 8.1e+03;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 EYLMKK 9
Db 2 ELSWQK 7

RESULT 11
O05991 PRELIMINARY; PRT; 15 AA.
AC O05991;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
```

RESULT 15			
Q9UMA0	PRELIMINARY;	PRT;	9 AA.
ID Q9UMA0			
AC Q9UMA0;			
DT 01-MAY-2000			
			(TReMBLrel. 13, Created)

DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE KIT protein (Fragment).
GN KIT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94061059; PubMed=7694728;
RA Spritz R.A., Holmes S.A., Berg S.Z., Nordlund J.J., Fukai K.;
RT "A recurrent deletion in the KIT (mast/stem cell growth factor
RL receptor) proto-oncogene is a frequent cause of human piebaldism."
RL Hum. Mol. Genet. 2:1499-1500(1993).
DR EMBL; S67686; AAD13996.1; -.
FT NON_TER 1
SQ SEQUENCE 9 AA; 1182 MW; 0BC504032361B5AB CRC64;

Query Match 23.88; Score 20; DB 4; Length 9;
Best Local Similarity 50.08; Pred. No. 6.7e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 7 WKQ 10
Db 5 WKRE 8

Search completed: January 28, 2003, 08:54:36
Job time : 32 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 08:42:09 ; Search time 34 Seconds
(without alignments)
62.706 Million cell updates/sec

Title: US-09-730-465-2_COPY_50_65
Perfect score: 84
Sequence: 1 PLKEVLWKQKQKVAE 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 255767

Minimum DB seq length: 0
Maximum DB seq length: 16

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*

- 1: /SID32/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SID32/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SID32/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SID32/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SID32/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SID32/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SID32/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SID32/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SID32/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SID32/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SID32/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SID32/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SID32/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SID32/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SID32/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SID32/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SID32/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SID32/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SID32/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SID32/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SID32/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SID32/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SID32/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					Description	
Result No.	Score	Query Match	Length	ID		
1	36	42.9	16	23	AAE23846	Human zsig33-gamma
2	36	42.9	16	23	AAE23847	Human zsig33-gamma
3	36	42.9	16	23	AAE15891	Human zsig33-gamma
4	36	42.9	16	23	AAE15892	Human zsig33-gamma
5	35	41.7	15	16	AAE73948	Fragment of T.serg
6	34	40.5	10	21	AAE69683	Human c-Erb-A nucl
7	34	40.5	15	23	AAE23848	Human zsig33-gamma
8	34	40.5	15	23	AAE15893	Human zsig33-gamma
9	32	38.1	14	22	AAE98523	Human peptide #179
10	32	38.1	16	15	AAE66027	Pair 3 Pep 5 immun

11	31	36.9	8	23	AAU76075	Nociceptin-like im
12	31	36.9	9	23	AAU76074	Nociceptin-like im
13	31	36.9	10	22	AAE31762	Amino acid sequenc
14	31	36.9	11	23	AAU76110	Nociceptin-like pe
15	31	36.9	12	23	AAU76099	Nociceptin-like pe
16	31	36.9	12	23	AAU76102	Nociceptin-like pe
17	31	36.9	12	23	AAU76106	Nociceptin-like pe
18	31	36.9	12	23	AAU76107	Nociceptin-like pe
19	31	36.9	12	23	AAU76112	Nociceptin-like pe
20	31	36.9	13	23	AAU76111	Nociceptin-like pe
21	31	36.9	14	18	AAE24446	Nucleic acid (NA)
22	31	36.9	14	22	AAE96920	Human peptide #195
23	31	36.9	14	22	AAE00706	Human protein frag
24	31	36.9	16	21	AAE52664	yes-encoded oncopr
25	30	35.7	8	18	AAE44604	Anti-fungal peptid
26	30	35.7	8	20	AAE00581	Anti-fungal peptid
27	30	35.7	8	22	AAE65505	Anti-fungal peptid
28	30	35.7	9	17	AAE04048	Anti-fungal peptid
29	30	35.7	9	18	AAE44562	Anti-fungal peptid
30	30	35.7	9	20	AAE00539	Anti-fungal peptid
31	30	35.7	9	22	AAE65463	Anti-fungal peptid
32	30	35.7	14	14	AAE31887	Protein S - C4BP b
33	30	35.7	16	15	AAE47955	CTLA-2 beta fragme
34	30	35.7	16	20	AAE42318	Synthetic cytotoxi
35	29	34.5	13	22	AAE03136	Human gene 7 encod
36	29	34.5	14	22	ABE56642	Human SNP related
37	29	34.5	15	16	AAE78029	BPI protein segmen
38	29	34.5	15	16	AAE81042	BPI.120 BPI domai
39	29	34.5	15	16	AAE87770	BPI.120 for use in
40	29	34.5	15	16	AAE76355	Bacterial permeabl
41	29	34.5	15	17	AAE05965	Recombinant BPI pe
42	29	34.5	15	18	AAE43657	Bactericidal/perme
43	29	34.5	15	19	AAE63416	Human BPI protein
44	29	34.5	15	21	AAE16155	Bactericidal/perme
45	29	34.5	15	22	AAE76846	Human colon cancer

ALIGNMENTS

RESULT 1
AAE23846
ID AAE23846 standard; peptide; 16 AA.
XX
AC AAE23846;
XX
DT 10-SEP-2002 (first entry)
XX
DE Human zsig33-gamma peptide #2.
XX
KW Human; zsig33-like peptide; gastric contractility; nutrient uptake;
KW growth hormone; digestive enzyme; restorative therapy; gene therapy;
KW protein therapy; gastrointestinal; endocrine; anabolic.
XX
OS Homo sapiens.
XX
PN US2002055156-A1.
XX
PD 09-MAY-2002.
XX
PF 10-MAY-2001; 2001US-0853253.
XX
PR 11-MAY-2000; 2000US-203300P.
XX
PA (JASP/) JASPERS S R.
PA (SHEP/) SHEPPARD P O.
PA (DEIS/) DEISHER T A.
PA (BISH/) BISHOP P D.
PI Jaspers SR, Sheppard PO, Deisher TA, Bishop PD;
XX WPI; 2002-443750/47.
XX

```

PT ZSIG33-Like peptides and polynucleotides, useful for modulating gastric
PT contractility, nutrient uptake, growth hormones and/or secretion of
PT digestive/pancreatic enzymes and hormones -
XX
XX Claim 5; Page 30; 34pp; English.
XX
XX The invention relates to zsig33-like peptides and their corresponding
XX nucleic acids and methods for modulating gastric contractility, nutrient
XX uptake, growth hormones, secretion of digestive enzymes and hormones.
XX The sequences of the invention are used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate ZSIG33 expression.
XX The nucleic acids of the invention and their complements are used as
XX DNA probes in diagnostic assays to detect and quantitate the presence
XX of similar nucleic acids in samples, and therefore which patients may be
XX in need of restorative therapy. The ZSIG33 peptides are used as antigens
XX in the production of antibodies against ZSIG33 and in assays to identify
XX modulators of ZSIG33 expression and activity. The anti-ZSIG33 antibodies
XX and antagonists are used to down regulate expression and activity. The
XX anti-ZSIG33 antibodies are also used as diagnostic agents for detecting
XX the presence of ZSIG33 in samples (e.g. by enzyme linked immunosorbent
XX assay (ELISA)). The peptides and nucleic acids of the invention are used
XX to modulate gastric contractility, nutrient uptake, growth hormones, the
XX secretion of digestive enzymes and hormones, and/or secretion of enzymes
XX and/or hormones in the pancreas. zsig33-like DNA is used in gene therapy
XX and zsig33-like peptide is used in protein therapy. The present sequence
XX is human zsig33-like peptide, zsig33-gamma peptide.
XX
SQ Sequence 16 AA;

Query Match 42.9%; Score 36; DB 23; Length 16;
Best Local Similarity 33.3%; Pred. No. 76;
Matches 5; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 2 LKEVLWKKQDKVAE 16
|:::|:::|:|:
Db 2 LQDILWEEAKEAPAD 16

RESULT 2
AAE23847
ID AAE23847 standard; peptide; 16 AA.
XX
XX AAE23847;
XX
XX 10-SEP-2002 (first entry)
XX
XX Human zsig33-gamma peptide #3.
XX
XX Human; zsig33-like peptide; gastric contractility; nutrient uptake;
XX growth hormone; digestive enzyme; restorative therapy; gene therapy;
XX protein therapy; gastrointestinal; endocrine; anabolic.
XX
XX Homo sapiens.
XX
XX US2002055156-A1.
XX
XX 09-MAY-2002.
XX
XX 10-MAY-2001; 2001US-0853253.
XX
XX 11-MAY-2000; 2000US-203300P.
XX
XX (JASP/) JASPERS S R.
XX (SHEP/) SHEPPARD P O.
XX (DEIS/) DEISHER T A.
XX (BISH/) BISHOP P D.
XX
XX Jaspers SR, Sheppard PO, Deisher TA, Bishop PD;
XX
XX WPI; 2002-443750/47.
XX
XX ZSIG33-Like peptides and polynucleotides, useful for modulating gastric
XX contractility, nutrient uptake, growth hormones and/or secretion of

```

```

PT digestive/pancreatic enzymes and hormones -
XX
XX Claim 5; Page 30; 34pp; English.
XX
XX The invention relates to zsig33-like peptides and their corresponding
XX nucleic acids and methods for modulating gastric contractility, nutrient
XX uptake, growth hormones, secretion of digestive enzymes and hormones.
XX The sequences of the invention are used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate ZSIG33 expression.
XX The nucleic acids of the invention and their complements are used as
XX DNA probes in diagnostic assays to detect and quantitate the presence
XX of similar nucleic acids in samples, and therefore which patients may be
XX in need of restorative therapy. The ZSIG33 peptides are used as antigens
XX in the production of antibodies against ZSIG33 and in assays to identify
XX modulators of ZSIG33 expression and activity. The anti-ZSIG33 antibodies
XX and antagonists are used to down regulate expression and activity. The
XX anti-ZSIG33 antibodies are also used as diagnostic agents for detecting
XX the presence of ZSIG33 in samples (e.g. by enzyme linked immunosorbent
XX assay (ELISA)). The peptides and nucleic acids of the invention are used
XX to modulate gastric contractility, nutrient uptake, growth hormones, the
XX secretion of digestive enzymes and hormones, and/or secretion of enzymes
XX and/or hormones in the pancreas. zsig33-like DNA is used in gene therapy
XX and zsig33-like peptide is used in protein therapy. The present sequence
XX is human zsig33-like peptide, zsig33-gamma peptide.
XX
SQ Sequence 16 AA;

Query Match 42.9%; Score 36; DB 23; Length 16;
Best Local Similarity 33.3%; Pred. No. 76;
Matches 5; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 2 LKEVLWKKQDKVAE 16
|:::|:::|:|:
Db 2 LQDILWEEAKEAPAD 16

RESULT 3
AAE15891
ID AAE15891 standard; peptide; 16 AA.
XX
XX AAE15891;
XX
XX 26-MAR-2002 (first entry)
XX
XX Human zsig33-gamma peptide #2.
XX
XX Human; zsig33-like peptide; ZS33LP; immunity; developmental process;
XX infection; human immunodeficiency virus; vaccine; antihypoglycaemic;
XX adsorption enhancer; gastrointestinal disease; growth related disease;
XX inflammation; gene therapy; growth regulation; blood vessel formation;
XX HIV; zsig33-gamma peptide.
XX
XX Homo sapiens.
XX
XX WO200187933-A2.
XX
XX 22-NOV-2001.
XX
XX 10-MAY-2001; 2001WO-US15091.
XX
XX 11-MAY-2000; 2000US-0569271.
XX
XX (ZYMO ) ZYMOGENETICS INC.
XX
XX Jaspers SR, Sheppard PO, Deisher TA, Bishop PD;
XX
XX WPI; 2002-082982/11.
XX N-PSDB; AAD25764.
XX
XX New polypeptides, useful for modulating gastric contractility, nutrient
XX uptake, pancreatic secretion of hormones, digestive enzymes and
XX treating gastrointestinal and growth related diseases, comprises
XX zsig33-like peptides -

```

XX PS Claim 5b; Page 85; 89pp; English.

XX CC The invention relates to zsig33-like peptides (ZS33LP) including

XX CC zsig33-linker, zsig33-beta, zsig33-gamma, zsig-33-delta and

XX CC zsig33-epsilon peptides and nucleic acid molecules encoding such

XX CC zsig33-like peptides. ZS33LP peptides activate the immune system

XX CC in boosting immunity to infectious diseases, treating immunocompromised

XX CC patients such as human immunodeficiency virus (HIV) patients, in

XX CC improving vaccines and in treatment of bacterial, viral, protozoal and

XX CC fungal infections. Peptides of the invention are used to identify and

XX CC isolate receptors involved in growth regulation in the liver, blood

XX CC vessel formation and other developmental processes. They are useful for

XX CC evaluating functions of hypothalamus-pituitary-adrenal axis, to modulate

XX CC growth and/or differentiation of tumour cells, as additives to anti-

XX CC hypoglycaemic preparations containing glucose and as adsorption

XX CC enhancers for oral drugs which require fast nutrient action and to

XX CC stimulate glucose-induced insulin release. They are also useful as

XX CC research reagents for the expansion, differentiation, growth factor and

XX CC hormone secretion and/or cell-cell interactions of tissues associated

XX CC with gastrointestinal system, brain and central nervous system. These

XX CC molecules are useful for treating dysfunction associated with contractile

XX CC tissues or to suppress or enhance contractility in vivo and to treat

XX CC gastrointestinal and growth related diseases. ZS33LP peptides, nucleic

XX CC acids and/or antibodies are useful for treating disorders associated

XX CC with gastrointestinal contractility, secretion of digestive enzymes,

XX CC hormone and acids, secretion of hormones in the pancreas and/or brain,

XX CC gastrointestinal motility, recruitment of digestive enzymes, inflammation

XX CC and regulation of nutrient absorption. Sequences of the invention are

XX CC useful in gene therapy. The present sequence is human zsig33-gamma

XX CC peptide.

XX SQ Sequence 16 AA;

Query Match 42.9%; Score 36; DB 23; Length 16;

Best Local Similarity 33.3%; Pred. No. 76;

Matches 5; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 2 LKEVLWKQDKVAE 16

1::|::|: |:

Db 2 LQDILWEAKEAPAD 16

RESULT 4

AAE15892

ID AAE15892 standard; peptide; 16 AA.

XX AC AAE15892;

XX DT 26-MAR-2002 (first entry)

XX DE Human zsig33-gamma peptide #3.

XX KW Human: zsig33-like peptide; ZS33LP; immunity; developmental process;

XX KW infection; human immunodeficiency virus; vaccine; antihypoglycaemic;

XX KW adsorption enhancer; gastrointestinal disease; growth related disease;

XX KW inflammation; gene therapy; growth regulation; blood vessel formation;

XX KW HIV; zsig33-gamma peptide.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Modified-site 16

FT /note= "C-terminal amide"

XX WO200187933-A2.

XX PD 22-NOV-2001.

XX PF 10-MAY-2001; 2001WO-US15091.

XX PR 11-MAY-2000; 2000US-0569271.

PA (ZYMO) ZYMOGENETICS INC.

XX Jaspers SR, Sheppard PO, Deisher TA, Bishop PD;

PI WPI: 2002-082982/11.

XX N-PSDB; AAD25764.

DR New polypeptides, useful for modulating gastric contractility, nutrient

XX uptake, pancreatic secretion of hormones, digestive enzymes and

PT treating gastrointestinal and growth related diseases, comprises

PT zsig33-like peptides -

XX Claim 5c; Page 85; 89pp; English.

XX The invention relates to zsig33-like peptides (ZS33LP) including

XX zsig33-linker, zsig33-beta, zsig33-gamma, zsig-33-delta and

XX zsig33-epsilon peptides and nucleic acid molecules encoding such

XX zsig33-like peptides. ZS33LP peptides activate the immune system

XX in boosting immunity to infectious diseases, treating immunocompromised

XX patients such as human immunodeficiency virus (HIV) patients, in

XX improving vaccines and in treatment of bacterial, viral, protozoal and

XX fungal infections. Peptides of the invention are used to identify and

XX isolate receptors involved in growth regulation in the liver, blood

XX vessel formation and other developmental processes. They are useful for

XX evaluating functions of hypothalamus-pituitary-adrenal axis, to modulate

XX growth and/or differentiation of tumour cells, as additives to anti-

XX hypoglycaemic preparations containing glucose and as adsorption

XX enhancers for oral drugs which require fast nutrient action and to

XX stimulate glucose-induced insulin release. They are also useful as

XX research reagents for the expansion, differentiation, growth factor and

XX hormone secretion and/or cell-cell interactions of tissues associated

XX with gastrointestinal system, brain and central nervous system. These

XX molecules are useful for treating dysfunction associated with contractile

XX tissues or to suppress or enhance contractility in vivo and to treat

XX gastrointestinal and growth related diseases. ZS33LP peptides, nucleic

XX acids and/or antibodies are useful for treating disorders associated

XX with gastrointestinal contractility, secretion of digestive enzymes,

XX hormone and acids, secretion of hormones in the pancreas and/or brain,

XX gastrointestinal motility, recruitment of digestive enzymes, inflammation

XX and regulation of nutrient absorption. Sequences of the invention are

XX useful in gene therapy. The present sequence is human zsig33-gamma

XX CC peptide.

XX SQ Sequence 16 AA;

Query Match 42.9%; Score 36; DB 23; Length 16;

Best Local Similarity 33.3%; Pred. No. 76;

Matches 5; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 2 LKEVLWKQDKVAE 16

1::|::|: |:

Db 2 LQDILWEAKEAPAD 16

RESULT 5

AAE15892

ID AAE15892 standard; Peptide; 15 AA.

XX AC AAE15892;

XX DT 04-DEC-1995 (first entry)

XX DE Fragment of T.sergenti 33 kDa protein.

XX KW bovine small piroplasma protozoa; 33 kDa protein; vaccine;

XX KW immunogen.

XX OS Theileria sergenti.

XX PN AU9470373-A.

XX PD 09-MAR-1995.

PF 22-AUG-1994; 94AU-0070373.
 PR 30-AUG-1993; 93JP-0238864.
 XX (NORQ) NAT INST ANIMAL HEALTH.
 PA (KAGA) CHERO SERO THERAPEUTIC RES INST.
 XX Fujisaki K, Kamio T, Kashiwazaki M, Kawazu S, Miyahara T;
 PI Nonaka F, Sakai E, Taneno A, Yamada S;
 XX WPI: 1995-147719/20.
 DR New bovine small piroplasma protozoa peptide(s) - contg. the
 PT Lys-Glu-Lys sequence, used for preventing infection or detecting
 PT antibodies.
 XX
 PS Claim 2: Page 16; 20pp; English.
 XX This sequence represents a fragment of the principal 33 kDa protein of
 CC T.sergenti (see also R73947). T.sergenti is a bovine small piroplasma
 CC protozoa. This sequence contains the KEK sequence which is the smallest
 CC immunogenic unit of the 33 kDa protein. This protein can be used in a
 CC vaccine for prevention of bovine small piroplasma disease.
 XX
 XX Sequence 15 AA;
 SQ
 Query Match 41.7%; Score 35; DB 16; Length 15;
 Best Local Similarity 55.6%; Pred. No. 1e+02;
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 Qy 4 EVLKKQKD 12
 Db 1 EVVWKEKE 9
 II::II::
 RESULT 6
 AAY69683
 ID AAY69683 standard; peptide: 10 AA.
 XX
 AC AAY69683;
 XX
 DT 08-MAY-2000 (first entry)
 DE Human c-Erb-A nuclear targeting sequence (NTS) C.
 XX
 KW Human; c-Erb-A; nuclear targeting sequence; NTS;
 KW gene-activated matrix; neuronal therapeutic agent; neurotrophic;
 KW neurotrophin family; FGF family; TGF-beta inhibitor; guidance factor;
 KW axonal generation; neurodegenerative disease; Alzheimer's disease;
 KW Parkinson's disease; Huntington's disease.
 XX
 OS Homo sapiens.
 XX
 PN WO966959-A2.
 XX
 XX 29-DEC-1999.
 PD
 XX
 XX 01-JUN-1999; 99WO-US12126.
 PF
 XX
 PR 01-JUN-1998; 98US-0088419.
 PR 23-OCT-1998; 98US-0178286.
 XX
 XX (SELE-) SELECTIVE GENETICS INC.
 PA
 XX
 XX Baird A, Berry M, Logan A, Gonzalez AM;
 PI WPI: 2000-160565/14.
 XX
 DR A gene-activated matrix device comprising a neuronal therapeutic
 PT encoding agent is capable of inducing neuronal axonal generation -
 XX Disclosure; Page 64; 125pp; English.
 PS
 XX

CC The invention relates to a novel device for promoting neuronal
 CC regeneration or altering neuronal growth, comprising a gene-activated
 CC matrix (GAM). The GAM comprises a biocompatible matrix, and at least one
 CC DNA encoding a neuronal therapeutic product operably linked to a
 CC promoter. The DNA is non-covalently associated with the matrix. The
 CC device may further comprise a targeting agent, which is complexed with
 CC or conjugated to DNA encoding a neuronal therapeutic agent, and which
 CC is capable of binding to a surface receptor of a neuronal or a repair
 CC cell. This targeting agent may also comprise a linker providing a
 CC nuclear or cytoplasmic translocation signal sequence. The
 CC invention also encompasses methods of preparing a GAM for promoting
 CC neuronal regeneration and survival. The neuronal therapeutic agent
 CC encoded by the GAM DNA is capable of inducing neuronal axonal generation
 CC or regeneration, and may include neurotrophic factors which are members
 CC of either the neurotrophin or FGF (fibroblast growth factor) families;
 CC inhibitors of antagonists of axonal generation (e.g., transforming
 CC growth factor (TGF)-beta inhibitors); or a guidance factor which
 CC promotes defined nerve growth. The device can be used to transfer the
 CC neuronal therapeutic-encoding DNA into a neuronal cell or repair cell.
 CC The device promotes axonal generation or regeneration without axonal
 CC entrapment. The device allows a neuronal connection to be established or
 CC re-established. The methods and devices of the invention are useful for
 CC treatment of neurodegenerative diseases such as Alzheimer's disease,
 CC Parkinson's disease or Huntington's disease. Direct plasmid DNA transfer
 CC from a matrix to a mammalian repair cell through stimulation of the
 CC wound healing process offers a number of advantages. Firstly, the ease
 CC of producing and purifying DNA constructs compares favourably with
 CC as structural scaffolds that promote cell ingrowth and proliferation.
 CC Secondly, direct gene transfer may be an advantageous method of
 CC delivery for molecules that normally undergo complex biosynthetic
 CC processing or for receptors, which must be properly positioned in the
 CC cell membrane. Sequences AAY69676-Y69693 represent nuclear targeting
 CC sequences (NTSs).
 XX
 XX Sequence 10 AA;
 SQ
 Query Match 40.5%; Score 34; DB 21; Length 10;
 Best Local Similarity 85.7%; Pred. No. 97;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 7 WKQKDK 13
 Db 3 WKQKDK 9
 IIIII
 RESULT 7
 AAE23848
 ID AAE23848 standard; peptide: 15 AA.
 XX
 AC AAE23848;
 XX
 DT 10-SEP-2002 (first entry)
 DE Human zsig33-gamma peptide #4.
 XX
 KW Human; zsig33-like peptide; gastric contractility; nutrient uptake;
 KW growth hormone; digestive enzyme; restorative therapy; gene therapy;
 KW protein therapy; gastrointestinal; endocrine; anabolic.
 XX
 OS Homo sapiens.
 XX
 PN US2002055156-A1.
 XX
 XX 09-MAY-2002.
 PD
 XX
 XX 10-MAY-2001; 2001US-0853253.
 PF
 XX
 PR 11-MAY-2000; 2000US-203300P.
 XX
 XX (JASP/) JASPERS S R.
 PA (SHEP/) SHEPPARD P O.
 PA (DEIS/) DEISHER T A.

OS Homo sapiens.
 PN WO200147944-A2.
 XX
 XX 05-JUL-2001.
 XX
 XX 28-DEC-2000; 2000WO-US35498.
 XX
 XX 28-DEC-1999; 99US-0173419.
 XX 27-DEC-2000; 2000US-0173419.
 XX
 XX (CURA-) CURAGEN CORP.
 XX
 XX Shimkets RA, Leach M;
 XX WPI; 2001-465210/50.
 XX
 XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
 PT oncogenes and histones, useful for diagnosing and treating, e.g.
 PT cancer, autoimmune diseases and infections -
 XX
 XX Disclosure; Page 4062; 4143pp; English.
 PS
 XX The present invention relates to oligonucleotides (see AAL26793-AAL34659)
 CC encoding polymorphic variants of proteins related to amylases, amyloid
 CC proteins, angiotensin, apoptosis related proteins, cadherin, cyclin,
 CC polymerase, oncogenes, histones, kinases, colony stimulating factors,
 CC complement related proteins, cytochromes, kinesins, cytokines,
 CC interferons, interleukins, G-protein coupled receptors and thioesterases.
 CC The present sequence is a peptide encoded by one such oligonucleotide.
 CC The oligonucleotides and the peptides encoded by them may be used in the
 CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate expression of the proteins listed above. Disorders that may
 CC be prevented, diagnosed and/or treated include multifactorial diseases
 CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid
 CC arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus
 CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
 CC brain, breast, colon and kidney, leukaemia), diseases of the nervous
 CC system and an infection of pathogenic organisms.
 XX
 XX Sequence 14 AA;
 SQ
 Query Match 38.1%; Score 32; DB 22; Length 14;
 Best Local Similarity 71.4%; Pred. No. 2.7e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 4 EYLWKKQ 10
 I :|||
 Db 6 ETWKKQ 12
 RESULT 10
 AAR66027
 ID AAR66027 standard; Protein; 16 AA.
 XX
 XX AAR66027;
 XX
 XX 08-JUN-1995 (first entry)
 XX
 XX Pair 3 Pep 5 immunogen for human PGMI isozyme allele.
 XX
 XX Phosphoglucomutase-1; PGMI; polymorphism; allele; antibody;
 XX immunogenic peptide.
 XX
 XX Synthetic.
 OS
 XX WO9424309-A.
 PN
 XX 27-OCT-1994.
 PD
 XX 18-APR-1994; 94WO-GB00809.
 XX
 XX 16-APR-1993; 93GB-0007895.
 PR

XX (MEDI-) MEDICAL RES COUNCIL.
 PA
 XX Edwards YH, Hopkinson DA, Whitehouse DB;
 PI
 XX WPI; 1994-341885/42.
 DR
 XX Test processes for detecting polymorphisms - useful for
 PT identifying alleles of human gene encoding phospho-glucomutase-1
 PT (PGMI)
 PT
 XX Disclosure; Table 5, page 12; 42pp; English.
 PS
 XX Human phosphoglucomutase-1 (PGMI) has four common alleles,
 CC designated PGMI*1+, 1-, 2+ and 2-. Two rare alleles, PGMI*3 and PGMI*
 CC 7 have also been identified. A point mutation in exon 4 at posn. 723
 CC from C to T correlates with Arg to Cys mutation at residue 220 of
 CC the protein and with the PGMI 1/2 protein polymorphism, whereas a
 CC point mutation in exon 8 at posn. 1320 from T to C leads to a Tyr-
 CC His substn. at residue 419 which shows complete association with the
 CC PGMI+/- protein polymorphism (see AAQ79276 & AAQ79277). There are also
 CC four less common variant isozymes, 3+, 3-, 7+ and 7-. The 3/7
 CC mutation is in exon 1A. Antibodies may be prepd. capable of
 CC distinguishing the 1/2 alleles or the +/- alleles using peptide
 CC immunogens based on fragment of the PGMI AA sequence incorporating
 CC one or other of the posn. 723 (see AAR66023/R66024) and/or posn.
 CC 1320 (see AAR66025/R66026) mutants. Peptides Pep 5 and Pep 6 (see
 CC AAR66017 & AAR66018) may similarly be used to generate allele-specific
 CC antibodies for the 3/7 polymorphism.
 XX
 SQ Sequence 16 AA;
 Query Match 38.1%; Score 32; DB 15; Length 16;
 Best Local Similarity 50.0%; Pred. No. 3e+02;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 2 LKEVLWKKQDK 13
 :|| :|||
 Db 4 MKEAIQKKKKK 15
 RESULT 11
 AAU76075
 ID AAU76075 standard; peptide; 8 AA.
 XX
 XX AC AAU76075;
 XX
 XX 08-MAY-2002 (first entry)
 DT
 XX Nociceptin-like immunogenic peptide #4.
 DE
 XX Nociceptin; opioid receptor-like 1; ORL1; hypotaemia;
 KW coronary heart failure; diuretic therapy; thiazide; loop diuretic;
 KW water diuresis; congestive heart failure; liver cirrhosis;
 KW nephrotic syndrome; hypertension; multiple organ failure;
 KW acute renal failure; hypokalaemia; oedema; antigen.
 XX
 OS Synthetic.
 OS
 XX WO200198324-A1.
 PN
 XX 27-DEC-2001.
 PD
 XX 15-JUN-2001; 2001WO-US19113.
 XX
 XX 16-JUN-2000; 2000DK-0000944.
 PR 05-OCT-2000; 2000DK-0001485.
 PR 06-DEC-2000; 2000US-251671P.
 PR 13-JUN-2001; 2001WO-US41008.
 PR
 XX (ZEAL-) ZEALAND PHARM AS.
 PA
 XX Larsen BD, Petersen JS, Kapusta DR, Harlow KW;
 PI

XX WPI; 2002-171551/22.

XX New peptide conjugate useful for preparing medicament for treating

PT congestive heart failure, liver cirrhosis, nephrotic syndrome and

PT hypertension comprises modified N and/or C terminals

XX Disclosure; Page 7; 72pp; English.

XX The invention relates to a peptide conjugate of the general formula (A).

CC R¹-L-X-Z'-R² (A); where X = a hexapeptide of formula (B);

CC R¹-L-X-Z'-R² (A); where X = a hexapeptide of formula (B);

CC A¹-A²-A³-A⁴-A⁵-A⁶ (B); A¹ = R, K, or H; A² = Y, W, or F; A³ = Y,

CC N, W or F; A⁴ = K, R or H; A⁵ = F, Y, W, L, V or I; and A⁶ = R, K or

CC H. Each amino acid residue in the hexapeptide may be in the L or D

CC form, Z and Z' = a charged peptide chain of 4-20 amino acid residues

CC having the D or L configuration or is missing provided that not both of Z

CC and Z' are missing; R¹ = H or an acyl group; and R² = NR³R⁴ or OH;

CC R³, R⁴ = H, C(1-6)alkoxy, arylloxy or a lower alkyl as defined,

CC where the conjugate being optionally further linked to a transport

CC moiety, and salts, hydrates and their solvates, and C-terminally amidated

CC or their esterified derivatives with suitable organic or inorganic acids.

CC Alternatively, the conjugate has a general formula (C): R¹-L-X-Z'-R² (C),

CC Where R¹, X, Z' and R² are same as defined in formula A; and salts,

CC hydrates and their solvates, and C-terminally amidated or their

CC esterified derivatives with suitable organic or inorganic acids.

CC The conjugate may also be linked to counterions selected from anions,

CC preferably CH₃COO⁻, CF₃COO⁻, Cl⁻, SO₃²⁻, maleate or oleate. Also

CC included are nucleic acids encoding the peptides, a host cell comprising/

CC expressing the peptides and antibodies against the peptides.

CC The peptides and conjugates are useful for the preparation of a

CC medicament for the treatment and/or prevention of hypotaemia which is

CC preferably associated with heart failure, or with intensive diuretic

CC therapy with thiazides and/or loop diuretics, water diuresis, congestive

CC heart failure, liver cirrhosis, nephrotic syndrome and hypertension,

CC multiple organ failure, acute renal failure, disease states associated

CC with elevated tone of nociceptin, hypokalaemia, oedema associated with

CC coronary heart failure. The hexapeptides are in part based on the

CC sequence of formula (R¹)(R²)(R³)(R⁴)(R⁵)(R⁶), a partial agonist of the

CC nociceptin, opioid receptor-like one (ORL1) which can be used to

CC raise antibodies against the conjugates. The present sequence is

CC an antigenic peptide of the invention.

XX Sequence 8 AA;

SQ Query Match 36.9%; Score 31; DB 23; Length 8;

Best Local Similarity 71.4%; Pred. No. 7.8e+05;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 WKQKDK 13

Db |||:|

2 WKKKKK 8

RESULT 12

AAU76074

ID AAU76074 standard; peptide; 9 AA.

XX AC AAU76074;

XX DT 08-MAY-2002 (first entry)

XX DE Nociceptin-like immunogenic peptide #3.

XX KW Nociceptin; opioid receptor-like 1; ORL1; hypotaemia;

KW coronary heart failure; diuretic therapy; thiazide; loop diuretic;

KW water diuresis; congestive heart failure; liver cirrhosis;

KW nephrotic syndrome; hypertension; multiple organ failure;

KW acute renal failure; hypokalaemia; oedema; antigen.

XX OS Synthetic.

XX PN WO200198324-A1.

XX

PD 27-DEC-2001.

XX 15-JUN-2001; 2001WO-US19113.

XX 16-JUN-2000; 2000DK-0000944.

PR 05-OCT-2000; 2000DK-0001485.

PR 06-DEC-2000; 2000US-251671P.

PR 13-JUN-2001; 2001WO-US41008.

XX (ZEAL-) ZEALAND PHARM AS.

XX Larsen BD, Petersen JS, Kapusta DR, Harlow KW;

XX WPI; 2002-171551/22.

XX New peptide conjugate useful for preparing medicament for treating

PT congestive heart failure, liver cirrhosis, nephrotic syndrome and

PT hypertension comprises modified N and/or C terminals

XX Disclosure; Page 7; 72pp; English.

XX The invention relates to a peptide conjugate of the general formula (A).

CC R¹-L-X-Z'-R² (A); where X = a hexapeptide of formula (B);

CC A¹-A²-A³-A⁴-A⁵-A⁶ (B); A¹ = R, K, or H; A² = Y, W, or F; A³ = Y,

CC N, W or F; A⁴ = K, R or H; A⁵ = F, Y, W, L, V or I; and A⁶ = R, K or

CC H. Each amino acid residue in the hexapeptide may be in the L or D

CC form, Z and Z' = a charged peptide chain of 4-20 amino acid residues

CC having the D or L configuration or is missing provided that not both of Z

CC and Z' are missing; R¹ = H or an acyl group; and R² = NR³R⁴ or OH;

CC R³, R⁴ = H, C(1-6)alkoxy, arylloxy or a lower alkyl as defined,

CC where the conjugate being optionally further linked to a transport

CC moiety, and salts, hydrates and their solvates, and C-terminally amidated

CC or their esterified derivatives with suitable organic or inorganic acids.

CC Alternatively, the conjugate has a general formula (C): R¹-L-X-Z'-R² (C),

CC Where R¹, X, Z' and R² are same as defined in formula A; and salts,

CC hydrates and their solvates, and C-terminally amidated or their

CC esterified derivatives with suitable organic or inorganic acids.

CC The conjugate may also be linked to counterions selected from anions,

CC preferably CH₃COO⁻, CF₃COO⁻, Cl⁻, SO₃²⁻, maleate or oleate. Also

CC included are nucleic acids encoding the peptides, a host cell comprising/

CC expressing the peptides and antibodies against the peptides.

CC The peptides and conjugates are useful for the preparation of a

CC medicament for the treatment and/or prevention of hypotaemia which is

CC preferably associated with heart failure, or with intensive diuretic

CC therapy with thiazides and/or loop diuretics, water diuresis, congestive

CC heart failure, liver cirrhosis, nephrotic syndrome and hypertension,

CC multiple organ failure, acute renal failure, disease states associated

CC with elevated tone of nociceptin, hypokalaemia, oedema associated with

CC coronary heart failure. The hexapeptides are in part based on the

CC sequence of formula (R¹)(R²)(R³)(R⁴)(R⁵)(R⁶), a partial agonist of the

CC nociceptin, opioid receptor-like one (ORL1) which can be used to

CC raise antibodies against the conjugates. The present sequence is

CC an antigenic peptide of the invention.

XX Sequence 9 AA;

SQ Query Match 36.9%; Score 31; DB 23; Length 9;

Best Local Similarity 71.4%; Pred. No. 7.8e+05;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 WKQKDK 13

Db |||:|

2 WKKKKK 8

RESULT 13

AAU76074

ID AAB31762 standard; peptide; 10 AA.

XX AC AAB31762;

XX DT 30-APR-2001 (first entry)

XX

DE Amino acid sequence of a cross-linking peptide.
 XX Nucleic acid condensate; cationic linker; gene therapy.
 KW Synthetic.
 XX WO200104135-A2.
 PN 18-JAN-2001.
 XX 13-JUL-2000; 2000WO-US19164.
 PF 13-JUL-1999; 99US-0143600.
 XX 05-OCT-1999; 99US-0157761.
 PR (UNMI) UNIV MICHIGAN.
 XX Rice KG, Adami RC, McKenzie DL, Collard WT, Kwok K, Park Y;
 PI Yang Y;
 PI WPI; 2001-168410/17.
 DR Compositions comprising nucleic acid condensates having a nucleic acid
 XX bound to two low molecular weight cationic linkers, used in human gene
 PT therapy, and diagnostics -
 PT Example 6; Page 108; 202pp; English.
 PS The specification describes a composition comprising a nucleic acid
 XX condensate. This condensate comprises a nucleic acid bound to two low
 CC molecular weight cationic linkers. The linkers are crosslinked to each
 CC other by reaction with a low molecular weight dialdehyde. Alternatively,
 CC the linkers each contain at least two thiol groups and are crosslinked
 CC by reaction of the thiol groups. The low molecular weight carriers are
 CC minimal in size, reduce toxicity, condense DNA into small particles,
 CC have increased stability, and mediate effective gene expression in a
 CC target tissue. The nucleic acid condensate is used for gene therapy,
 CC particularly human gene therapy, and diagnostics. It is also used for
 CC expressing nucleic acids in cells and providing a nucleic acid to an
 CC animal. The present sequence represents a cross-linking peptide, which
 CC is used as a linker in the composition of the invention. The peptide
 CC condenses DNA.
 XX Sequence 10 AA;
 SQ Query Match 36.9%; Score 31; DB 22; Length 10;
 Best Local Similarity 71.4%; Pred. No. 2.7e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 7 WKKQDK 13
 Db |||:|
 2 WKKKKK 8
 RESULT 14
 AAU76110
 ID AAU76110 standard; peptide; 11 AA.
 XX AAU76110;
 AC 08-MAY-2002 (first entry)
 XX Nociceptin-like peptide conjugate compound 12.
 DE Nociceptin; opioid receptor-like 1; ORL1; hypotaemia;
 XX coronary heart failure; diuretic therapy; thiazide; loop diuretic;
 KW water diuresis; congestive heart failure; liver cirrhosis;
 KW nephrotic syndrome; hypertension; multiple organ failure;
 KW acute renal failure; hypokalaemia; oedema.
 XX Synthetic.
 OS Key Location/Qualifiers
 FH

FT Modified-site 1 /note= "Arg is acetylated"
 FT Modified-site 11 /note= "Lys is amidated"
 FT WO200198324-A1.
 PN 27-DEC-2001.
 XX 15-JUN-2001; 2001WO-US19113.
 PF 16-JUN-2000; 2000DK-0000944.
 XX 05-OCT-2000; 2000DK-0001485.
 PR 06-DEC-2000; 2000US-251671P.
 XX 13-JUN-2001; 2001WO-US41008.
 XX (ZEAL-) ZEALAND PHARM AS.
 PA Larsen BD, Petersen JS, Kapusta DR, Harlow KW;
 PI WPI; 2002-171551/22.
 DR New peptide conjugate useful for preparing medicament for treating
 XX congestive heart failure, liver cirrhosis, nephrotic syndrome and
 PT hypertension comprises modified N and/or C terminals -
 PT Claim 19; Page 52; 72pp; English.
 PS The invention relates to a peptide conjugate of the general formula (A).
 XX R₁-Z-X-Z'-R₂ (A); where X = a hexapeptide of formula (B);
 CC A¹-A²-A³-A⁴-A⁵-A⁶ (B); A¹ = R, K, or H; A² = Y, W, or F; A³ = Y,
 CC N, W or F; A⁴ = K, R or H; A⁵ = F, Y, W, L, V or I; and A⁶ = R, K or
 CC H. Each amino acid residue in the hexapeptide may be in the L or D
 CC form. Z and Z' = a charged peptide chain of 4-20 amino acid residues
 CC having the D or L configuration or is missing provided that not both of Z
 CC and Z' are missing; R₁ = H or an acyl group; and R₂ = Np³⁸*4 or OH;
 CC R³, R⁴ = H, C(1-6)alkoxy, aryloxy or a lower alkyl as defined,
 CC where the conjugate being optionally further linked to a transport
 CC moiety, and salts, hydrates and their solvates, and C-terminally amidated
 CC or their esterified derivatives with suitable organic or inorganic acids.
 CC Alternatively, the conjugate has a general formula (C). R₁-X-Z'-R₂ (C),
 CC where R₁, X, Z' and R₂ are same as defined in formula A; and salts,
 CC hydrates and their solvates, and C-terminally amidated or their
 CC esterified derivatives with suitable organic or inorganic acids.
 CC The conjugate may also be linked to counterions selected from anions,
 CC preferably CH₃COO⁻, CF₃COO⁻, Cl⁻, SO₃*2⁻, maleate or oleate. Also
 CC included are nucleic acids encoding the peptides, a host cell comprising/
 CC expressing the peptides and antibodies against the peptides.
 CC The peptides and conjugates are useful for the preparation of a
 CC medicament for the treatment and/or prevention of hypotaemia which is
 CC preferably associated with heart failure, or with intensive diuretic
 CC therapy with thiazides and/or loop diuretics, water diuresis, congestive
 CC heart failure, liver cirrhosis, nephrotic syndrome and hypertension,
 CC multiple organ failure, acute renal failure, disease states associated
 CC with elevated tone of nociceptin, hypokalaemia, oedema associated with
 CC coronary heart failure. The hexapeptides are in part based on the
 CC sequence of formula (RK)YY(RK)(WI)(RK), a partial agonist of the
 CC nociceptin, opioid receptor-like one (ORL1) which can be used to raise
 CC antibodies against the conjugates. The present sequence is
 CC a peptide conjugate of the invention.
 XX Sequence 11 AA;
 SQ

Query Match 36.9%; Score 31; DB 23; Length 11;
 Best Local Similarity 71.4%; Pred. No. 3e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 7 WKKQDK 13
 |||:|

Db 5 WKKKKK 11

RESULT 15

AAU76099
ID AAU76099 standard; peptide; 12 AA.
AC AAU76099;
XX 08-MAY-2002 (first entry)
XX Nociceptin-like peptide conjugate compound 1.
XX
XX Nociceptin; opioid receptor-like 1; ORL1; hypotaemia;
KW coronary heart failure; diuretic therapy; thiazide; loop diuretic;
KW water diuresis; congestive heart failure; liver cirrhosis;
KW nephrotic syndrome; hypertension; multiple organ failure;
KW acute renal failure; hypokalaemia; oedema.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 1 /note= "Arg is acetylated"
FT Modified-site 12 /note= "Lys is amidated"
FT
XX WO200198324-A1.
XX
XX 27-DEC-2001.
XX
XX 15-JUN-2001; 2001WO-US19113.
XX
XX 16-JUN-2000; 2000DK-0000944.
PR 05-OCT-2000; 2000DK-0001485.
PR 06-DEC-2000; 2000US-251671P.
PR 13-JUN-2001; 2001WO-US41008.
XX
XX (ZEAL-) ZEALAND PHARM AS.
XX
XX Larsen BD, Petersen JS, Kapusta DR, Harlow KW;
XX WPI; 2002-171551/22.
XX
XX New peptide conjugate useful for preparing medicament for treating
PT congestive heart failure, liver cirrhosis, nephrotic syndrome and
PT hypertension comprises modified N and/or C terminals
XX
XX Claim 19; Page 52; 72pp; English.
XX
XX The invention relates to a peptide conjugate of the general formula (A).
CC R₁-Z-X'-R₂ (A); where X = a hexapeptide of formula (B);
CC A¹-A²-A³-A⁴-A⁵-A⁶ (B); A¹ = R, K, or H; A² = Y, W, or F; A³ = Y,
CC N, W or F; A⁴ = K, R or H; A⁵ = F, Y, W, L, V or I; and A⁶ = R, K or
CC H. Each amino acid residue in the hexapeptide may be in the L or D
CC form, Z and Z' = a charged peptide chain of 4-20 amino acid residues
CC having the D or L configuration or is missing provided that not both of Z
CC and Z' are missing; R¹ = H or an acyl group; and R² = NR³R⁴ or OH;
CC R³, R⁴ = H, C(1-6)alkoxy, aryloxy or a lower alkyl as defined,
CC where the conjugate being optionally further linked to a transport
CC moiety, and salts, hydrates and their solvates, and C-terminally amidated
CC or their esterified derivatives with suitable organic or inorganic acids.
CC Alternatively, the conjugate has a general formula (C). R₁-X'-Z'-R₂ (C).
CC Where R₁, X, Z' and R₂ are same as defined in formula A; and salts,
CC hydrates and solvates, and C-terminally amidated or their
CC esterified derivatives with suitable organic or inorganic acids.
CC The conjugate may also be linked to counterions selected from anions,
CC preferably CH₃COO⁻, CF₃COO⁻, Cl⁻, SO₃²⁻, maleate or oleate. Also
CC included are nucleic acids encoding the peptides, a host cell comprising/
CC expressing the peptides and antibodies against the peptides.
CC The peptides and conjugates are useful for the preparation of a
CC medicament for the treatment and/or prevention of hypotaemia which is
CC preferably associated with heart failure, or with intensive diuretic
CC therapy with thiazides and/or loop diuretics, water diuresis, congestive
CC heart failure, liver cirrhosis, nephrotic syndrome and hypertension,
CC multiple organ failure, acute renal failure, disease states associated
CC with elevated tone of nociceptin, hypokalaemia, oedema associated with

CC coronary heart failure. The hexapeptides are in part based on the
CC sequence of formula (RK)YY(RK)(WL)(RK), a partial agonist of the
CC nociceptin, opioid receptor-like one (ORL1) which can be used to raise
CC antibodies against the conjugates. The present sequence is
XX a peptide conjugate of the invention.
SQ Sequence 12 AA;
Query Match 36.9%; Score 31; DB 23; Length 12;
Best Local Similarity 71.4%; Pred. No. 3.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 7 WKQKQDK 13
||:|
Db 5 WKKKKKK 11
Search completed: January 28, 2003, 08:53:42
Job time : 36 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 08:54:40 ; Search time 11 Seconds
(without alignments)
29.351 Million cell updates/sec

Title: US-09-730-465-2_COPY_50_65

Perfect score: 84

Sequence: 1 PLKEVLWKKQDKVAE 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 31315

Minimum DB seq length: 0

Maximum DB seq length: 16

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	42.9	16	10	US-09-853-253-15
2	36	42.9	16	10	US-09-853-253-16
3	34	40.5	15	10	US-09-853-253-17
4	30	35.7	9	10	US-09-881-490-163
5	29	34.5	15	10	US-09-765-527-101
6	26	31.0	9	10	US-09-071-838-74
7	26	31.0	10	10	US-09-984-056-85
8	26	31.0	10	10	US-09-984-057-85
9	26	31.0	11	10	US-09-791-378-302
10	26	31.0	14	10	US-09-766-396-8
11	26	31.0	14	12	US-10-062-375-8
12	26	31.0	15	10	US-09-766-396-23
13	26	31.0	15	10	US-09-813-653-1
14	26	31.0	15	12	US-10-062-375-23
15	26	31.0	16	9	US-09-981-286A-4
16	26	31.0	16	10	US-09-854-204-58
17	26	31.0	16	10	US-09-756-983-3
18	25	29.8	7	10	US-09-854-204-9
19	25	29.8	8	10	US-09-765-527-122

Sequence 3, Appli
Sequence 26, Appl
Sequence 48, Appl
Sequence 4, Appli
Sequence 8, Appli
Sequence 124, App
Sequence 27, Appl
Sequence 44, Appl
Sequence 198, App
Sequence 199, App
Sequence 5, Appli
Sequence 6, Appli
Sequence 125, App
Sequence 7, Appli
Sequence 8, Appli
Sequence 94, Appl
Sequence 123, App
Sequence 36, Appl
Sequence 9, Appli
Sequence 2, Appli
Sequence 16, Appl
Sequence 17, Appl
Sequence 18, Appl
Sequence 19, Appl
Sequence 20, Appl

20 25 29.8 9 9 US-09-805-301-3
21 25 29.8 9 10 US-09-124-280A-26
22 25 29.8 9 10 US-09-872-832-48
23 25 29.8 10 9 US-09-805-301-4
24 25 29.8 10 10 US-09-765-527-8
25 25 29.8 10 10 US-09-765-527-124
26 25 29.8 10 10 US-09-124-280A-27
27 25 29.8 10 10 US-09-124-280A-44
28 25 29.8 10 10 US-09-881-490-198
29 25 29.8 10 10 US-09-881-490-199
30 25 29.8 11 9 US-09-805-301-5
31 25 29.8 12 9 US-09-805-301-6
32 25 29.8 12 10 US-09-765-527-125
33 25 29.8 13 9 US-09-805-301-7
34 25 29.8 13 9 US-09-805-301-8
35 25 29.8 14 10 US-09-765-527-94
36 25 29.8 14 10 US-09-765-527-123
37 25 29.8 14 10 US-09-881-490-36
38 25 29.8 15 9 US-09-805-301-9
39 25 29.8 15 10 US-09-765-527-2
40 25 29.8 15 10 US-09-765-527-9
41 25 29.8 15 10 US-09-765-527-16
42 25 29.8 15 10 US-09-765-527-17
43 25 29.8 15 10 US-09-765-527-18
44 25 29.8 15 10 US-09-765-527-19
45 25 29.8 15 10 US-09-765-527-20

ALIGNMENTS

RESULT 1
US-09-853-253-15
; Sequence 15, Application US/09853253
; Patent No. US20020055156A1
; GENERAL INFORMATION:
; APPLICANT: JASPERS, STEPHEN
; APPLICANT: SHEPPARD, PAUL
; APPLICANT: DEISHER, THERESA
; APPLICANT: BISHOP, PAUL
; TITLE OF INVENTION: Zsig33-like Peptides
; FILE REFERENCE: 00-30
; CURRENT APPLICATION NUMBER: US/09/853,253
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: 60/203,300
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-853-253-15

Query Match 42.9%; Score 36; DB 10; Length 15;
Best Local Similarity 33.3%; Pred. No. 8.1;
Matches 5; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 2 LKEVLWKKQDKVAE 16
|::|::|::|:
Db 2 LQDILWEEKEAPAD 16

RESULT 2
US-09-853-253-16
; Sequence 16, Application US/09853253
; Patent No. US20020055156A1
; GENERAL INFORMATION:
; APPLICANT: JASPERS, STEPHEN
; APPLICANT: SHEPPARD, PAUL
; APPLICANT: DEISHER, THERESA
; APPLICANT: BISHOP, PAUL
; TITLE OF INVENTION: Zsig33-like Peptides

```
; FILE REFERENCE: 00-30
; CURRENT APPLICATION NUMBER: US/09/853,253
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: 60/203,300
; PRIOR FILING DATE: 2000-05-11
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: AMIDATION
; LOCATION: (16)...(16)
US-09-853-253-16

Query Match 42.9%; Score 36; DB 10; Length 16;
Best Local Similarity 33.3%; Pred. No. 8.1;
Matches 5; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 2 LKEVLWKQKDKVAE 16
Db 2 LQDILWEEAKEAPAD 16

RESULT 3
US-09-853-253-17
; Sequence 17, Application US/09853253
; Patent No. US20020055156A1
; GENERAL INFORMATION:
; APPLICANT: JASPERS, STEPHEN
; APPLICANT: SHEPPARD, PAUL
; APPLICANT: DEISHER, THERESA
; APPLICANT: BISHOP, PAUL
; TITLE OF INVENTION: Zsig33-like Peptides
; FILE REFERENCE: 00-30
; CURRENT APPLICATION NUMBER: US/09/853,253
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: 60/203,300
; PRIOR FILING DATE: 2000-05-11
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: AMIDATION
; LOCATION: (15)...(15)
US-09-853-253-17

Query Match 40.5%; Score 34; DB 10; Length 15;
Best Local Similarity 35.7%; Pred. No. 15;
Matches 5; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 2 LKEVLWKQKDKVA 15
Db 2 LQDILWEEAKEAPA 15

RESULT 4
US-09-881-490-163
; Sequence 163, Application US/09881490
; Patent No. US20020077298A1
; GENERAL INFORMATION:
; APPLICANT: Little II, Roger G.
; Lim, Edward
; Fadem, Mitchell B.
; TITLE OF INVENTION: Anti-Fungal Peptides
; NUMBER OF SEQUENCES: 211
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th FloorDrive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/881,490
; FILING DATE: 14-Jun-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/119,858
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 08/372,105
; FILING DATE: 13-JAN-95
; APPLICATION NUMBER: 08/306,473
; FILING DATE: 15-SEP-94
; APPLICATION NUMBER: 08/273,540
; FILING DATE: 11-JUL-94
; APPLICATION NUMBER: 08/209,762
; FILING DATE: 11-MAR-94
; APPLICATION NUMBER: 08/183,222
; FILING DATE: 14-JAN-94
; APPLICATION NUMBER: 08/093,202
; FILING DATE: 15-JUL-93
; APPLICATION NUMBER: 08/030,644
; FILING DATE: 12-MAR-93
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 100-238/11021US01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/707-8889
; TELEFAX: 312/707-9155
; TELEX: 650 388-1248
; INFORMATION FOR SEQ ID NO: 163:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: "XMP.332"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1-9
; OTHER INFORMATION: /label= D-Amino Acids
; /note= "Positions 1-9 are D-amino acids"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: C-Terminus
; OTHER INFORMATION: /label= Amidation
; /note= "The C-Terminus is Amidated"
; SEQUENCE DESCRIPTION: SEQ ID NO: 163:
US-09-881-490-163

Query Match 35.7%; Score 30; DB 10; Length 9;
Best Local Similarity 66.7%; Pred. No. 1e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 EVLWKK 9
Db 4 QILWKK 9

RESULT 5
US-09-765-527-101
; Sequence 101, Application US/09765527
; Patent No. US2002006638A1
; GENERAL INFORMATION:
```

APPLICANT: Better, Marc D.
TITLE OF INVENTION: Methods for Recombinant Microbial Production of
Fusion Proteins and BPI-Derived Peptides
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,527
FILING DATE: 18-Jan-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/621,803
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Borun, Michael F.
REGISTRATION NUMBER: 25,447
REFERENCE/DOCKET NUMBER: 27129/33199
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 101:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: "XMP.120"
FEATURE:
NAME/KEY: Modified-site
LOCATION: C-Terminus
OTHER INFORMATION: /label= Amidation
/note= "The C-Terminus is Amidated."
SEQUENCE DESCRIPTION: SEQ ID NO: 101:
US-09-765-527-101
Query Match 34.5%; Score 29; DB 10; Length 15;
Best Local Similarity 62.5%; Pred. No. 88;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 7 WKQKQDKV 14
Db 7 WRAQRKRL 14
RESULT 6
US-09-071-838-74
Sequence 74, Application US/09071838
Patent No. US20020152501a1
GENERAL INFORMATION:
APPLICANT: Fischer, Robert L.
APPLICANT: Ohad, Nir
APPLICANT: Kiyosue, Tomohiro
APPLICANT: Yadegari, Ramin
APPLICANT: Margossian, Linda
APPLICANT: Harada, John
APPLICANT: Goldberg, Robert B.
TITLE OF INVENTION: Nucleic Acids That Control Seed and
Fruit Development in Plants
NUMBER OF SEQUENCES: 324
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,838
FILING DATE: 01-MAY-1998
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-086100US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 74:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-071-838-74
Query Match 31.0%; Score 26; DB 10; Length 9;
Best Local Similarity 50.0%; Pred. No. 1e+05;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy 6 LWKQKDK 13
Db 2 IWEKQVK 9
RESULT 7
US-09-984-056-85
Sequence 85, Application US/09984056
Patent No. US20020120106A1
GENERAL INFORMATION:
APPLICANT: BOGOCH, SAMUEL
APPLICANT: BOGOCH, ELEONORE S.
TITLE OF INVENTION: ANTHRAX AND SMALL POX REPLICONS AND METHODS OF USE
FILE REFERENCE: 09425-46903
CURRENT APPLICATION NUMBER: US/09/984,056
CURRENT FILING DATE: 2001-10-26
PRIOR APPLICATION NUMBER: 60/303,396
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: 60/278,761
PRIOR FILING DATE: 2001-03-27
PRIOR APPLICATION NUMBER: 09/146,755
PRIOR FILING DATE: 1998-09-04
PRIOR APPLICATION NUMBER: 09/817,144
PRIOR FILING DATE: 2001-03-27
PRIOR APPLICATION NUMBER: 08/198,139
PRIOR FILING DATE: 1994-02-17
NUMBER OF SEQ ID NOS: 103
SOFTWARE: PatentIn 2.1
SEQ ID NO 85
LENGTH: 10
TYPE: PRT
ORGANISM: Homo sapiens
US-09-984-056-85
Query Match 31.0%; Score 26; DB 10; Length 10;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 8 KKQKDK 13
|||

```
Db      5 KKHDK 10

RESULT 8
US-09-984-057-85
; Sequence 85, Application US/09984057
; Patent No. US20020151677A1
; GENERAL INFORMATION:
; APPLICANT: BOGOCH, SAMUEL
; APPLICANT: BOGOCH, ELENORE S.
; TITLE OF INVENTION: REPLIKIN-CONTAINING SEQUENCES
; FILE REFERENCE: 09425-46902
; CURRENT APPLICATION NUMBER: US/09/984,057
; CURRENT FILING DATE: 2001-10-26
; PRIOR FILING DATE: 2001-10-26
; PRIOR FILING DATE: 2001-07-09
; PRIOR FILING DATE: 2001-07-09
; PRIOR FILING DATE: 2001-03-27
; PRIOR FILING DATE: 1998-09-04
; PRIOR FILING DATE: 2001-03-27
; PRIOR FILING DATE: 2001-03-27
; PRIOR FILING DATE: 1994-02-17
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn 2.1
; SEQ ID NO 85
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-057-85

Query Match      31.0%; Score 26; DB 10; Length 10;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      8 KQKDK 13
      |||||
Db      5 KKHDK 10

RESULT 9
US-09-791-378-302
; Sequence 302, Application US/09791378
; Patent No. US20020142303A1
; GENERAL INFORMATION:
; APPLICANT: Patekh, Rajesh
; TITLE OF INVENTION: PROTEINS, GENES AND THEIR USE FOR DIAGNOSIS AND TREATMENT OF
; FILE REFERENCE: 9195-061-999
; CURRENT APPLICATION NUMBER: US/09/791,378
; CURRENT FILING DATE: 2001-02-23
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 677
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 302
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-378-302

Query Match      31.0%; Score 26; DB 10; Length 11;
Best Local Similarity 55.6%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      5 VLKWKQDK 13
      |||||
Db      3 VEWQLQDK 11

RESULT 10
US-09-766-396-8
; Sequence 8, Application US/09766396
; Patent No. US20020013456A1
; GENERAL INFORMATION:
; APPLICANT: Sutcliffe, Gregor J.
; APPLICANT: de Lecea, Luis
; APPLICANT: Henriksen, Steven J.
; TITLE OF INVENTION: NEUROPEPTIDES,
; COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
; STREET: 10666 No. US20020013456Alth Torrey Pines Road, TPC-8
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/766,396
; FILING DATE: 18-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/857,389
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Schmonsees, William
; REGISTRATION NUMBER: 31,796
; REFERENCE/DOCKET NUMBER: 22908-0002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-7041
; TELEFAX: (415) 324-0638
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: C-terminal
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-766-396-8

Query Match      31.0%; Score 26; DB 10; Length 14;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 PLKEVLWK 8
      |||||
Db      1 PCKNFWK 8

RESULT 11
US-10-062-375-8
; Sequence 8, Application US/10062375
; Patent No. US20020133000A1
; GENERAL INFORMATION:
; APPLICANT: Sutcliffe, Gregor J.
; APPLICANT: de Lecea, Luis
; APPLICANT: Siggs, George R.
; APPLICANT: Henriksen, Steven J.
; TITLE OF INVENTION: NEUROPEPTIDES,
; COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
; STREET: 10666 No. US20020133000Alth Torrey Pines Road, TPC-8
; CITY: La Jolla
; STATE: California
```


NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE
STREET: 10666 NO. US20020133000Alch Torrey Pines Road, TPC-8
CITY: La Jolla
STATE: California
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/062.375
FILING DATE: 30-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/857.389
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Schmonsees, William
REGISTRATION NUMBER: 31,796
REFERENCE/DOCKET NUMBER: 22908-0002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-7041
TELEFAX: (415) 324-0638
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: C-terminal
SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-10-062-375-23

Query Match 31.0%; Score 26; DB 12; Length 15;
Best Local Similarity 50.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 PLKEVLWK 8
| | | |
Db 2 PCKNEFWK 9

RESULT 15
US-09-981-286A-4
Sequence 4, Application US/09981286A
Publication No. US20020192799A1
GENERAL INFORMATION:
APPLICANT: Watowich, Stanley J.
APPLICANT: Weaver, Scott C.
APPLICANT: Davey, Robert A.
TITLE OF INVENTION: Drug Discovery Methods
FILE REFERENCE: 265.00260101
CURRENT APPLICATION NUMBER: US/09/981.286A
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: US 60/240,187
PRIOR FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 16
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Cell-permeant polypeptide
US-09-981-286A-4

Query Match 31.0%; Score 26; DB 9; Length 16;
Best Local Similarity 44.4%; Pred. No. 2.7e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 PLKEVLWK 9
| : : | | |
Db 8 PNRMMKWK 16

Search completed: January 28, 2003, 08:58:42
Job time : 11 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 08:53:06 ; Search time 34 seconds
(without alignments)
13.846 Million cell updates/sec

Title: US-09-730-465-2_COPY_50_65

Perfect score: 84

Sequence: 1 PLKEVLWKKOKDKVAE 16

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 116189

Minimum DB seq length: 0

Maximum DB seq length: 16

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	50.0	7	6	Patent No. 5185441
2	31	36.9	10	4	Sequence 5, Appli
3	31	36.9	15	4	Sequence 6, Appli
4	31	36.9	15	4	Sequence 11, Appl
5	30	35.7	8	2	Sequence 205, App
6	30	35.7	9	2	Sequence 163, App
7	30	35.7	9	5	Sequence 13, Appl
8	30	35.7	14	1	Sequence 11, Appl
9	30	35.7	16	2	Sequence 11, Appl
10	30	35.7	16	2	Sequence 11, Appl
11	30	35.7	16	5	Sequence 116, App
12	29	34.5	15	1	Sequence 116, App
13	29	34.5	15	1	Sequence 116, App
14	29	34.5	15	1	Sequence 116, App
15	29	34.5	15	1	Sequence 116, App
16	29	34.5	15	1	Sequence 116, App
17	29	34.5	15	2	Sequence 101, App
18	29	34.5	15	3	Sequence 116, App
19	29	34.5	15	4	Sequence 116, App
20	29	34.5	15	4	Sequence 116, App
21	29	34.5	15	4	Sequence 101, App
22	29	34.5	15	5	Sequence 116, App
23	29	34.5	15	5	Sequence 116, App
24	28	33.3	11	4	Sequence 1, Appli
25	28	33.3	11	4	Sequence 23, Appl
26	28	33.3	15	1	Sequence 17, Appl
27	28	33.3	16	4	Sequence 32, Appl

28 27.5 32.7 12 2 US-08-659-984A-8 Sequence 8, Appli
29 27.5 32.7 12 4 US-08-660-531-8 Sequence 6, Appli
30 27.5 32.7 14 2 US-08-659-984A-11 Sequence 11, Appl
31 27.5 32.7 14 4 US-08-660-531-11 Sequence 11, Appl
32 27 32.1 5 1 US-07-657-769B-36 Sequence 36, Appl
33 27 32.1 5 1 US-07-789-184-97 Sequence 37, Appl
34 27 32.1 5 1 US-08-475-263-97 Sequence 97, Appl
35 27 32.1 5 1 US-08-485-886-97 Sequence 97, Appl
36 27 32.1 5 2 US-08-477-362-97 Sequence 97, Appl
37 27 32.1 5 2 US-08-477-134-97 Sequence 97, Appl
38 27 32.1 5 3 US-08-473-489A-97 Sequence 97, Appl
39 27 32.1 5 3 US-08-485-695-97 Sequence 97, Appl
40 27 32.1 5 4 US-08-018-760-97 Sequence 97, Appl
41 27 32.1 10 1 US-08-347-000-7 Sequence 7, Appli
42 27 32.1 10 1 US-08-378-761A-46 Sequence 46, Appl
43 27 32.1 10 1 US-08-485-286-46 Sequence 46, Appl
44 27 32.1 10 6 5248606-35 Patent No. 5248606
45 27 32.1 11 3 US-08-970-833-8 Sequence 8, Appli

ALIGNMENTS

RESULT 1
5185441-12
; Patent No. 5185441
; APPLICANT: WALLNER, BARBARA P.; HESSIONS, CATHERINE
; TITLE OF INVENTION: DNA SEQUENCES, RECOMBINANT DNA
; MOLECULES AND PROCESSES FOR PRODUCING PI-LINKED LYMPHOCYTE
; FUNCTION ASSOCIATED ANTIGEN-3
; NUMBER OF SEQUENCES: 41
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/237,309
; FILING DATE: 26-AUG-1988
; SEQ ID NO:12:
; LENGTH: 7
5185441-12

Query Match 50.0%; Score 42; DB 6; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.9e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WKKOKK 13
Db 1 WKKOKK 7

RESULT 2
US-09-050-811-5
; Sequence 5, Application US/09050811
; Patent No. 6387700
; GENERAL INFORMATION:
; APPLICANT: Rice, Kevin G.
; APPLICANT: Wadhwa, Manpreet S.
; TITLE OF INVENTION: Peptides for Gene Delivery
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/050,811
; FILING DATE: 30-MAR-1998
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:

```
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: UM-03314
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/705-8410
; TELEFAX: 415/397-8338
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Binding-site
; LOCATION: 1
; OTHER INFORMATION: /note= "The residue at this
; position is bound to an Alkaloid by a Sulfide."
US-09-050-811-5

Query Match 36.9%; Score 31; DB 4; Length 10;
Best Local Similarity 71.4%; Pred. No. 64;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 7 WKQKDK 13
   |||:|
Db 2 WKKKKK 8

RESULT 3
US-09-050-811-6
; Sequence 6, Application US/09050811
; Patent No. 6387700
; GENERAL INFORMATION:
; APPLICANT: Rice, Kevin G.
; APPLICANT: Wadhwa, Manpreet S.
; TITLE OF INVENTION: Peptides for Gene Delivery
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/050,811
; FILING DATE: 30-MAR-1998
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: UM-03314
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/705-8410
; TELEFAX: 415/397-8338
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Binding-site
; LOCATION: 1
; OTHER INFORMATION: /note= "The residue at this
; position is bound to an Alkaloid by a Sulfide."
```

```
US-09-050-811-6

Query Match 36.9%; Score 31; DB 4; Length 15;
Best Local Similarity 71.4%; Pred. No. 95;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 7 WKQKDK 13
   |||:|
Db 2 WKKKKK 8

RESULT 4
US-09-050-811-11
; Sequence 11, Application US/09050811
; Patent No. 6387700
; GENERAL INFORMATION:
; APPLICANT: Rice, Kevin G.
; APPLICANT: Wadhwa, Manpreet S.
; TITLE OF INVENTION: Peptides for Gene Delivery
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/050,811
; FILING DATE: 30-MAR-1998
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: UM-03314
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/705-8410
; TELEFAX: 415/397-8338
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-050-811-11

Query Match 36.9%; Score 31; DB 4; Length 15;
Best Local Similarity 71.4%; Pred. No. 95;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 7 WKQKDK 13
   |||:|
Db 2 WKKKKK 8

RESULT 5
US-08-621-259A-205
; Sequence 205, Application US/08621259A
; Patent No. 5858974
; GENERAL INFORMATION:
; APPLICANT: Little II, Roger G
; APPLICANT: Lim, Edward
; APPLICANT: Fadem, Mitchell B.
; TITLE OF INVENTION: Anti-Fungal Peptides
; NUMBER OF SEQUENCES: 252
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
```

```

; STREET: 500 West Madison Street
; City: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/621,259A
; FILING DATE: 21-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/504,841
; FILING DATE: 20-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 11021US02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/707-8889
; TELEFAX: 312/707-9155
; TELEX:
; INFORMATION FOR SEQ ID NO: 205:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: "XMP.374"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1-8
; OTHER INFORMATION: /label= D-Amino Acids
; OTHER INFORMATION: /note= "Positions 1-8 are D-Amino Acids."
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1-8
; OTHER INFORMATION: /label= D-Amino Acids
; OTHER INFORMATION: /note= "Positions 1-8 are D-Amino Acids."
; US-08-621-259A-205

```

```

Query Match 35.7%; Score 30; DB 2; Length 8;
Best Local Similarity 66.7%; Pred. No. 1.9e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 4 EVLWKK 9
;:::|
Db 3 QILWKK 8

```

```

RESULT 6
US-08-621-259A-163
; Sequence 163, Application US/08621259A
; Patent No. 5858974
; GENERAL INFORMATION:
; APPLICANT: Little II, Roger G
; APPLICANT: Lim, Edward
; TITLE OF INVENTION: Anti-Fungal Peptides
; NUMBER OF SEQUENCES: 252
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/621,259A
; FILING DATE: 21-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/504,841
; FILING DATE: 20-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 11021US02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/707-8889
; TELEFAX: 312/707-9155
; TELEX:
; INFORMATION FOR SEQ ID NO: 163:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: "XMP.332"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1-9
; OTHER INFORMATION: /label= D-Amino Acids
; OTHER INFORMATION: /note= "Positions 1-9 are D-amino acids"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: C-Terminus
; OTHER INFORMATION: /label= Amidation
; OTHER INFORMATION: /note= "The C-Terminus is Amidated."
; US-08-621-259A-163

```

```

Query Match 35.7%; Score 30; DB 2; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.9e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 4 EVLWKK 9
;:::|
Db 4 QILWKK 9

```

```

RESULT 7
PCT-US95-09262-163
; Sequence 163, Application PC/TUS9509262
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Anti-Fungal Peptides
; NUMBER OF SEQUENCES: 206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09262
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/372,105
; FILING DATE: 13-JAN-95
; PRIOR APPLICATION DATA:

```

```
; APPLICATION NUMBER: 08/306,473
; FILING DATE: 15-SEP-94
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/273,540
; FILING DATE: 11-JUL-94
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/209,762
; FILING DATE: 11-MAR-94
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/183,222
; FILING DATE: 14-JAN-94
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/093,202
; FILING DATE: 15-JUL-93
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/030,644
; FILING DATE: 12-MAR-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25,447
; REFERENCE/DOCKET NUMBER: 27129/10040
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 163:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: "XMP.332"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1-9
; OTHER INFORMATION: /label= D-Amino Acids
; OTHER INFORMATION: /note= "Positions 1-9 are D-amino acids"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: C-Terminus
; OTHER INFORMATION: /label= Amidation
; OTHER INFORMATION: /note= "The C-Terminus is Amidated"
; PCT-US95-09262-163

Query Match 35.7%; Score 30; DB 5; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.9e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 EVLWKK 9
Db 4 QILWKK 9

RESULT 8
US-07-907-190-13
; Sequence 13, Application US/07907190
; Patent No. 5321123
; GENERAL INFORMATION:
; APPLICANT: Griffin, John H.
; APPLICANT: Fernandez, Jose A.
; TITLE OF INVENTION: PROTEIN S POLYPEPTIDES AND ANTI-PEPTIDE
; TITLE OF INVENTION: ANTIBODIES THAT INHIBIT PROTEIN S BINDING TO C4b BINDING
; TITLE OF INVENTION: PROTEIN, DIAGNOSTIC SYSTEMS AND THERAPEUTIC METHODS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5321123th Torrey Pines Road, Mail Drop TPC8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
```

```
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/907,190
; FILING DATE: 19920701
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/724,746
; FILING DATE: 02-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SC11119P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-07-907-190-13

Query Match 35.7%; Score 30; DB 1; Length 14;
Best Local Similarity 55.6%; Pred. No. 1.3e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 LKEVLWKKQ 10
Db 3 INELIQKKQ 11

RESULT 9
US-08-485-937-11
; Sequence 11, Application US/08485937
; Patent No. 5962633
; GENERAL INFORMATION:
; APPLICANT: Muller, Daniel
; APPLICANT: Brownell, Elise
; APPLICANT: Delaria, Katherine
; APPLICANT: Wallace, Linda
; TITLE OF INVENTION: Cytotoxic T-Lymphocyte Antigen as a
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Allegretti, Ltd.
; STREET: 10 S. Wacker Drive Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,937
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,215
; FILING DATE: 18-MAY-1995
; APPLICATION NUMBER: WO PCT/US93/06552
; FILING DATE: 15-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/915,923
; FILING DATE: 17-JUL-1992
```


APPLICANT: Lambert, Jr., Lewis H.
TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial
Infection by Administration of
Bactericidal/Permeability-Increasing
TITLE OF INVENTION:
NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/311,611A
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/273,401
FILING DATE: 11-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/125,651
FILING DATE: 22-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Sharp, Jeffrey S.
REGISTRATION NUMBER: 31,879
REFERENCE/DOCKET NUMBER: 32251
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 116:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: "BPI.120"
US-08-311-611A-116

Query Match 34.5%; Score 29; DB 1; Length 15;
Best Local Similarity 62.5%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 WKKQKDKV 14
||| |
DB 7 WKAQKRKL 14

RESULT 13
US-08-372-783-116
Sequence 116, Application US/08372783
Patent No. 5578572
GENERAL INFORMATION:
APPLICANT: Horwitz, Arnold H.
APPLICANT: Lambert, Lewis H.
APPLICANT: Little, Roger G.
TITLE OF INVENTION: Anti-gram-Positive Bacterial Methods and
Materials
NUMBER OF SEQUENCES: 237
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606-6402

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/372,783
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/273,540
FILING DATE: 11-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/209,762
FILING DATE: 11-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/183,222
FILING DATE: 14-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27129/32415
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 116:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: "XMP.120"
US-08-372-783-116

Query Match 34.5%; Score 29; DB 1; Length 15;
Best Local Similarity 62.5%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 WKKQKDKV 14
||| |
DB 7 WKAQKRKL 14

RESULT 14
US-08-372-105-116
Sequence 116, Application US/08372105
Patent No. 5627153
GENERAL INFORMATION:
APPLICANT: Little, Roger G.
APPLICANT: Lim, Edward
APPLICANT: Lambert, Lewis H.
APPLICANT: Scannon, Patrick J.
TITLE OF INVENTION: Anti-Fungal Materials and Methods
NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/372,105
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/273,540

```
; FILING DATE: 11-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/209,762
; FILING DATE: 11-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/183,222
; FILING DATE: 14-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27129/32415
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 116:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: "XMP.120"
US-08-372-105-116

Query Match 34.5%; Score 29; DB 1; Length 15;
Best Local Similarity 62.5%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 7 WKQKDKV 14
   |||||
Db 7 WKAQKRKL 14
```

```
RESULT 15
US-08-306-473A-116
; Sequence 116, Application US/08306473A
; Patent No. 5652332
; GENERAL INFORMATION:
; APPLICANT: Little, Roger G.
; TITLE OF INVENTION: Biologically Active Peptides from
; TITLE OF INVENTION: Functional Domains of Bactericidal/
; TITLE OF INVENTION: Permeability-Increasing Protein and
; TITLE OF INVENTION: Uses Thereof
; NUMBER OF SEQUENCES: 226
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: Suite 3000, 10 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/306,473A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/209,762
; FILING DATE: 11-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/183,222
; FILING DATE: 14-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J.
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 93,1133-
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
```

```
; TELEFAX: 312-715-1234
; INFORMATION FOR SEQ ID NO: 116:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: "BPI.120"
US-08-306-473A-116

Query Match 34.5%; Score 29; DB 1; Length 15;
Best Local Similarity 62.5%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 7 WKQKDKV 14
   |||||
Db 7 WKAQKRKL 14
```

```
Search completed: January 28, 2003, 08:55:39
Job time : 36 secs
```

10

RESULT 3

E49757

techoic acid synthesis protein tagC' - *Bacillus subtilis* (strain 168)C:Species: *Bacillus subtilis*

C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 15-Oct-1999

C:Accession: E49757

R:MaueI, C.; Young, M.; Karamata, D.

J. Gen. Microbiol. 137, 929-941, 1991

A:Title: Genes concerned with synthesis of poly(glycerol phosphate), the essential teichoic acid synthesis protein tagC' - *Bacillus subtilis* (strain 168)

A:Reference number: A49757; MUID:913111391; PMID:1306926

A:Accession: E49757

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-40 <MAU>

A:Cross-references: GB:M57497; NID:g143722; PIDN:AAA22846.1; PID:g143726

Query Match 12.8%; Score 42; DB 2; Length 40;

Best Local Similarity 66.7%; Pred. No. 2.3e+02;

Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 CFGFISCFSQOI 14

|| ||| |||

DB 22 CFNVISCIWQOI 33

RESULT 4

H81170

conserved hypothetical protein NMB0674 [imported] - *Neisseria meningitidis* (strain MC58)C:Species: *Neisseria meningitidis*

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 02-Feb-2001

C:Accession: H81170

R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.

Li, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.

Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V.

A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.

A:Reference number: A81000; MUID:20175755; PMID:10710307

A:Accession: H81170

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-60 <TET>

A:Cross-references: GB:AE002422; GB:AE002098; NID:g7225898; PIDN:ANF41092.1; PID:g722590

A:Experimental source: serogroup B, strain MC58

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel

; Holroyd, S.; Jagers, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,

Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.

A:Reference number: A81775; MUID:20222556; PMID:10761919

A:Accession: E81933

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-60 <PAR>

A:Cross-references: GB:AL162754; GB:AL157959; NID:g7379424; PIDN:CAB84154.1; PID:g737958

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: NMB0674; NNA0874

Query Match

Best Local Similarity 27.7%; Score 42; DB 2; Length 60;

Matches 13; Conservative 7; Mismatches 15; Indels 12; Gaps 2;

QY 17 VYGNVTFHVPSPVPLKEVLWKQ-----KDKVAELENSEAFPS 56

| : : : | : | : | : | : | : |

DB 13 VTKGRLEYHQD-----KQELNSRQAKLAYPIKDGIPYMLENEARALS 54

RESULT 5

B81350

small hydrophobic protein Cj0786 [imported] - *Campylobacter jejuni* (strain NCTC 11168)C:Species: *Campylobacter jejuni*

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002

C:Accession: B81350

R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chil

C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Bar

Nature 403, 665-668, 2000

A:Title: The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals

A:Reference number: A81250; MUID:20150912; PMID:10688204

A:Accession: B81350

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-57 <PAR>

A:Cross-references: GB:AL139076; GB:AL111168; NID:g6969128; PIDN:CAB73051.1; PID:g696

A:Experimental source: serotype O2, strain NCTC 11168

C:Genetics:

A:Gene: Cj0786

Query Match 12.7%; Score 41.5; DB 2; Length 57;

Best Local Similarity 31.6%; Pred. No. 4e+02;

Matches 12; Conservative 9; Mismatches 16; Indels 1; Gaps 1;

QY 14 IYGVYGNVTFHVPSPVPLKEVLWKVKOKKVAELENSE 51

| : : : | : | : | : | : | : |

DB 10 IIGVIFG-IDYVYFKNQDLKVQVKKQKELNSSLKEK 46

RESULT 6

JQ0254

hypothetical 4K protein (psbH-petB intergenic region) - rice chloroplast

C:Species: chloroplast *Oryza sativa* (rice)

C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 21-Jul-2000

C:Accession: JQ0254; S05134

R:Shimada, H.; Whittier, R.F.; Hiratsuka, J.; Maeda, Y.; Hirai, A.; Sugiyara, M.

submitted to JIPID, December 1989

A:Reference number: JQ0200

A:Accession: JQ0254

A:Molecule type: DNA

A:Residues: 1-35 <SHI>

A:Experimental source: cv. Nihonbare

R:Hiratsuka, J.; Shimada, H.; Whittier, R.; Ishibashi, T.; Sakamoto, M.; Morl, M.; Ko

Mol. Gen. Genet. 217, 185-194, 1989

A:Title: The complete sequence of the rice (*Oryza sativa*) chloroplast genome: intermo

of the cereals.

A:Reference number: S05080; MUID:89364698; PMID:2770692

A:Accession: S05134

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-35 <HIR>

A:Cross-references: GB:X15901; NID:g11957; PIDN:CAA33974.1; PID:g12014

A:Experimental source: cv. Nihonbare

A>Note: this sequence was submitted to EMBL, July 1989

C:Genetics:

A:Genome: chloroplast

C:Keywords: chloroplast; transmembrane protein

Query Match

Best Local Similarity 35.0%; Score 41; DB 2; Length 35;

Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 16 GVVYGNVTFHVPSPVPLKEV 35

| : : : | : | : | : |

DB 15 GIIFFAIFREPPKVPKVV 34

A:Reference number: JC6187; MUID:97169156; PMID:9016963

A:Accession: JC6188

A:Molecule type: DNA

A:Residues: 1-55 <SHC>

A:Cross-references: EMBL:X92859

C:Comment: This protein is a transcription factor to act in the forespore chamber of the

C:Genetics:

A:Gene: CSGAB

A:Start codon: TTG

Query Match 12.5%; Score 41; DB 2; Length 55;

Best Local Similarity 61.5%; Pred. No. 4.4e+02;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 27 PSNVPLKEVLWK 39

Db 30 PSMCTIKENLWK 42

RESULT 8

C75475

hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000

C:Accession: C75475

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Otterback, T.; Zaleski, C.; Ma

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: C75475

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-60 <WHI>

A:Cross-references: GB:AE001934; GB:AE000513; NID:96458504; PIDN:AAF10380.1; PID:9645851

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR0800

A:Map position: 1

Query Match 12.5%; Score 41; DB 2; Length 60;

Best Local Similarity 35.7%; Pred. No. 4.8e+02;

Matches 10; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

QY 23 TFHVPSNVPLKEVLWKQKDKVAELEN 50

Db 25 TGNVDANGMDKMSDKERADKLSELDN 52

RESULT 9

RHPGCE

corticoliberin - pig

N:Alternate names: corticotropin-releasing factor; CRF; endorpholiberin

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 11-Apr-1997

C:Accession: A01404

R:Patthy, M.; Horvath, J.; Mason-Garcia, M.; Szoke, B.; Schlesinger, D.H.; Schally, A.V.

Proc. Natl. Acad. Sci. U.S.A. 82, 8762-8766, 1985

A:Title: Isolation and amino acid sequence of corticotropin-releasing factor from pig hy

A:Reference number: A01404; MUID:86094305; PMID:3878520

A:Accession: A01404

A:Molecule type: protein

A:Residues: 1-41 <PAT>

A>Note: 40-Ile was also found

C:Comment: This hormone from hypothalamus regulates the release of corticotropin from pi

C:Superfamily: corticoliberin-endorpholiberin; diuretic hormone homology

C:Keywords: amidated carboxyl end; hormone; hypothalamus

F:41/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 12.4%; Score 40.5; DB 1; Length 41;

Best Local Similarity 26.3%; Pred. No. 3.6e+02;

Matches 10; Conservative 12; Mismatches 11; Indels 5; Gaps 1;

QY 21 NVTFHVPSNVPLKEVLWKQKDKVAELENSEPRAFSS 58

Db 9 DLTFHL-----LREVLAMARAEOLAQAHSNRKLMENF 41

RESULT 10

PW0053

angiotensin converting enzyme like protein (EC 3.4.21.1) - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 10-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 03-Nov-2000

C:Accession: PW0053

R:Sugimura, K.; Tian, X.L.; Hoffmann, S.; Ganten, D.; Bader, M.

Biochem. Biophys. Res. Commun. 247, 466-472, 1998

A:Title: Alternative splicing of the mRNA coding for the human endothelial angiotensi

A:Reference number: PW0053; MUID:98308142; PMID:9642152

A:Accession: PW0053

A:Molecule type: mRNA

A:Residues: 1-31 <SUC>

C:Superfamily: mammalian peptidyl-dipeptidase A

C:Keywords: hydrolase

Query Match 12.2%; Score 40; DB 2; Length 31;

Best Local Similarity 75.0%; Pred. No. 3e+02;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 24 FHVPSNVP 31

Db 4 FHIPSSVP 11

RESULT 11

AB1924

photosystem II protein [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp.

A:Note: Nostoc sp. Strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002

C:Accession: AB1924

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AB1924

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-39 <KUR>

A:Cross-references: GB:BA000019; PIDN:BAW72898.1; PID:gl7130287; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: psbX

Query Match 12.1%; Score 39.5; DB 2; Length 39;

Best Local Similarity 26.5%; Pred. No. 4.5e+02;

Matches 9; Conservative 10; Mismatches 12; Indels 3; Gaps 1;

QY 11 SQQIGVGVYGNVTFHVPSNVPLKEVLWKQKDKV 44

Db 6 SNFLSLAWGLTIIVIPATVGL---IFISQKDKI 36

RESULT 12

B97505

hypothetical protein AGR_C_2198 [imported] - Agrobacterium tumefaciens (strain C58, C

C:Species: Agrobacterium tumefaciens

C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002

C:Accession: B97505

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorillo, B.; Goldm

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz,

Science 294, 2323-2328, 2001

A:Title: Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium

A:Reference number: A97359; PMID:11743194

A:Accession: B97505

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: January 28, 2003, 08:55:45 ; Search time 10 Seconds

(without alignments)
253.006 Million cell updates/sec

Title: US-09-730-465-2_COPY_20_80

Perfect score: 327

Sequence: 1 LHCFGFISCFQIYGVYGVG.....DKVAELENSEFRASFKNR 61

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 5368

Minimum DB seq length: 0

Maximum DB seq length: 61

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	44	13.5	22	1 CR31_LITSP	P56238 litoria spl
2	41	12.5	35	1 PSBT_ORYSA	P12183 oryza sativ
3	40	12.2	22	1 CR32_LITCE	P56239 litoria cae
4	39.5	12.1	52	1 YAAB_BACSU	P37525 bacillus su
5	39	11.9	57	1 RK32_ODOSI	P49564 odontella s
6	38.5	11.8	39	1 PSBX_CYAPA	P48266 cyanophora
7	38	11.6	57	1 IVBC_NAJNA	P19859 naja naja (
8	37	11.3	22	1 CR33_LITCE	P56240 litoria cae
9	37	11.3	33	1 PSBT_MAIZE	P37257 zea mays (m
10	37	11.3	34	1 PSBT_TOBAC	P12184 nicotiana t
11	37	11.3	49	1 YS02_BORBU	O54556 borrelia bu
12	37	11.3	49	1 YU04_BORBU	O51014 borrelia bu
13	37	11.3	55	1 SYV_MACFA	P36391 macaca fasc
14	37	11.3	58	1 FAT_HV1B5	P04612 human immun
15	36	11.0	31	1 SODC_STRHE	P81163 striga herm
16	36	11.0	35	1 TPA_THADA	P21381 thaumatococ
17	36	11.0	52	1 NU3M_NEUCR	Q35141 neurospora
18	36	11.0	56	1 RL33_RICCN	Q92fw7 rickettsia
19	36	11.0	56	1 YOR7_TTV1	P19282 thermoprote
20	35.5	10.9	53	1 NCE1_YEAST	Q02820 saccharomyc
21	35.5	10.9	61	1 RL35_MVCFE	Q05428 mycoplasma
22	35	10.7	35	1 PSBT_MARPO	P12182 marchantia
23	35	10.7	36	1 Y260_BACHD	O8K953 bacillus ha
24	35	10.7	56	1 RL33_RICPR	Q92c89 rickettsia
25	34.5	10.6	55	1 RLX_HALMA	P14125 haloarcula
26	34	10.4	22	1 CR34_LITCE	P56241 litoria cae
27	34	10.4	40	1 PRE_BACLI	P18189 bacillus li
28	34	10.4	46	1 DEQO_BACLI	P12051 bacillus li
29	34	10.4	57	1 V3A_IBVU5	P30240 avian infec
30	34	10.4	60	1 RS21_MVCPN	P57079 mycoplasma
31	33.5	10.2	30	1 CALM_LYTP1	P05935 lytechninus
32	33.5	10.2	51	1 ATP8_LIMPO	Q9mlq5 limulus pol
33	33.5	10.2	59	1 R17E_HALN1	Q9hpX1 halobacteri

34	33.5	10.2	59	1 YF6A_HAEIN	O85243 haemophilus
35	33	10.1	35	1 PSBT_PINTH	P41625 pinus thunb
36	33	10.1	36	1 MFG2_DAGCL	Q41183 dactylis gl
37	33	10.1	37	1 DIU1_TENMO	P56618 tenebrio mo
38	33	10.1	40	1 CSP2_STRMT	O33666 streptococ
39	33	10.1	43	1 DHBD_ASPOR	P80402 aspergillus
40	33	10.1	54	1 Y01K_BP74	P39228 bacterioph
41	33	10.1	58	1 V07K_STMV	P17575 satellite t
42	32.5	9.9	55	1 YHFG_ECOLI	P37770 escherichia
43	32.5	9.9	59	1 RL35_MYCCE	P47439 mycoplasma
44	32.5	9.9	60	1 YH14_ARCFU	O28559 archaeoglob
45	32	9.8	23	1 CR42_LITCE	P56243 litoria cae

ALIGNMENTS

RESULT 1					
CR31_LITSP					
ID	CR31_LITSP	STANDARD;	PRT;	22 AA.	
AC	P56238;				
DT	15-JUL-1998 (Rel. 36, Created)				
DT	15-JUL-1998 (Rel. 36, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Caerin 3.1.				
OS	Litoria splendida (Magnificent tree frog), and				
OS	Litoria gilleni (Centralian tree frog).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;				
OC	Pelodyadinae; Litoria.				
OX	NCBI_TaxID=30345, 39405;				
RN	[1]				
RP	SEQUENCE, AND MASS SPECTROMETRY.				
RC	SPECIES=L splendida; TISSUE=Parotoid gland;				
RA	Stone D.J.M., Waugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.;				
RT	"Peptides from Australian frogs. Structures of the caerins and				
RT	caeridin 1 from Litoria splendida.";				
RL	J. Chem. Soc. Perkin Trans. 1:3173-3178(1992).				
[2]					
RP	SEQUENCE, AND MASS SPECTROMETRY.				
RC	SPECIES=L.gilleni; TISSUE=Parotoid gland;				
RA	Waugh R.J., Stone D.J.M., Bowie J.H., Wallace J.C., Tyler M.J.;				
RT	"Peptides from Australian frogs. The structures of the caerins and				
RT	caeridins from Litoria gilleni.";				
RL	J. Chem. Res. 139:937-961(1993).				
CC	-!- FUNCTION: ANTIBACTERIAL PEPTIDE, THAT ADOPTS AN ALPHA HELICAL				
CC	CONFORMATION WHICH CAN DISRUPT BACTERIAL MEMBRANES. EACH CAERIN				
CC	DISPLAYS A DIFFERENT ANTIMICROBIAL SPECIFICITY.				
CC	-!- TISSUE SPECIFICITY: SECRETED BY THE SKIN PAROTOID AND/OR ROSTRAL				
CC	GLANDS.				
CC	-!- MASS SPECTROMETRY: MW=2382; METHOD=FAB.				
KW	Antibiotic; Amphibian skin; Amidation.				
FT	MOD_RES 22				
SQ	SEQUENCE 22 AA; 2385 MW; 1D4411E2E9D43739 CRC64;				
Query Match 13.5%; Score 44; DB 1; Length 22;					
Best Local Similarity 66.7%; Pred. No. 41;					
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;					
OY 36 LMKKOKDKVAEL 47					
: : : :					
Db 2 LMQKDKKASEL 13					
RESULT 2					
PSBT_ORYSA					
ID	PSBT_ORYSA	STANDARD;	PRT;	35 AA.	
AC	P12183;				
DT	01-OCT-1989 (Rel. 12, Created)				
DT	01-OCT-1989 (Rel. 12, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Photosystem II reaction center T protein.				
GN	PSBT OR YCF8.				

```

RA Stone D.J.M., Waugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.;
RT "Peptides from Australian frogs. The structures of the caerins from
RL Litoria caerulea.";
RC J. Chem. Res. 138:910-936(1993).
RR
CC -!- FUNCTION: ANTIBACTERIAL PEPTIDE, THAT ADOPTS AN ALPHA HELICAL
CC CONFORMATION WHICH CAN DISRUPT BACTERIAL MEMBRANES. EACH CABRIN
CC DISPLAYS A DIFFERENT ANTIMICROBIAL SPECIFICITY.
CC
CC -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN PAROTOID AND/OR ROSTRAL
CC GLANDS.
CC
CC -!- MASS SPECTROMETRY: MW=2397; METHOD=FAB.
CC Antibiocytic; Amphibian skin; Amidation.
KW MOD_RES 22 22
FT AMIDATION.
SQ SEQUENCE 22 AA; 2400 MW; 1D40B3829D4367C CRC64;

Query Match 12.2%; Score 40; DB 1; Length 22;
Best Local Similarity 58.3%; Pred. NO. 1.2e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps

QY 36 LKKKKDKVAEL 47
DB 2 LWEKKEKASEL 13
1111111111

RESULT 4
YAAB_BACSU STANDARD; PRT; 52 AA.
ID YAAB_BACSU
AC P37525;
AD 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DT Hypothetical protein yaab.
GN YAAB
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
[1]
RX SEQUENCE FROM N.A.
RY MEDLINE=45215612; PubMed=2987847;
RR Moriya S., Ogasawara N., Yoshikawa H.;
RT "structure and function of the region of the replication origin of
RL the Bacillus subtilis chromosome. III. Nucleotide sequence of some
RT 10,000 base pairs in the origin region.";
RL Nucleic Acids Res. 13:2251-2265(1985).
[2]
RX SEQUENCE FROM N.A.
RY STRAIN=168;
RC MEDLINE=96051385; PubMed=7584024;
RR Ogasawara N., Nakai S., Yoshikawa H.;
RT "Systematic sequencing of the 180 kilobase region of the Bacillus
RL subtilis chromosome containing the replication origin.";
RT DNA Res. 1:1-14(1994).
[3]
RX SEQUENCE FROM N.A.
RY STRAIN=168;
RC MEDLINE=98044033; PubMed=9384377;
RR Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Boltin A., Borchert S.,
RA Borris R., Boursier L.L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschii C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devigne K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauei C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogilwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,

```


DE Venom chymotrypsin inhibitor.
OS Naja naja (Indian cobra).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Elapinae; Naja.
OX NCBI_TaxID=35670;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=91085579; PubMed=2262001;
RA Shafqat J., Zaidi Z.H., Joernvall H.;
RT "Purification and characterization of a chymotrypsin Kunitz inhibitor
RL type of polypeptide from the venom of cobra (Naja naja naja).";
FEBS Lett. 275:6-8(1990).
CC -!- FUNCTION: THIS PROTEIN INHIBITS CHYMOTRYPSIN.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: CONTAINS 1 BPTI/KUNITZ INHIBITOR DOMAIN.
DR PIR: S12957; S12957.
DR HSSP: P00981; 1DTK.
DR InterPro: IPR002223; Kunitz_BPTI.
DR Pfam: PF00014; Kunitz_BPTI; 1.
DR PRINTS: PR00759; BASICTPTASE.
DR ProDom: PD000222; Kunitz_BPTI; 1.
DR SMART: SM00131; KU; 1.
DR PROSITE: PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE: PS0279; BPTI_KUNITZ_2; 1.
KW Venom; Serine protease inhibitor.
FT DISULFID 5 55 BY SIMILARITY.
FT DISULFID 14 38 BY SIMILARITY.
FT DISULFID 30 51 BY SIMILARITY.
FT ACT_SITE 15 16 REACTIVE BOND (BY SIMILARITY).
SQ SEQUENCE 57 AA; 6508 MW; 4BFB26A010C1737A CRC64;

Query Match 11.6%; Score 38; DB 1; Length 57;
Best Local Similarity 31.88; Pred. No. 5.5e+02;
Matches 7; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 3 CFGTSCFSQQIYGVGVGVNTF 24
|||||:|:|:|:
DB 14 CFGFVSSYYNRYNSCHTSFTY 35

RESULT 8
CR33_LITCE STANDARD; PRT; 22 AA.
AC P56240;
DT 13-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Caerin 3.3.
OS Litoria caerulea (Green tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Pelodyadinae; Litoria.
OX NCBI_TaxID=30344;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=Parotoid gland;
RX Stone D.J.M., Waugh R.J., Bowie J.H., Wallace J.C., Tyler M.J.;
RA "Peptides from Australian frogs. The structures of the caerins from
RT Litoria caerulea";
RL J. Chem. Res. 136:910-936(1993).
CC -!- FUNCTION: ANTIBACTERIAL PEPTIDE, THAT ADOPTS AN ALPHA HELICAL
CC CONFORMATION WHICH CAN DISRUPT BACTERIAL MEMBRANES. EACH CAERIN
CC DISPLAYS A DIFFERENT ANTIMICROBIAL SPECIFICITY.
CC -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN PAROTOID AND/OR ROSTRAL
CC GLANDS.
CC -!- MASS SPECTROMETRY: MW=2424; METHOD=FAB.
KW Antibiotic; Amphibian skin; Amidation.
FT MOD_RES 22 22 AMIDATION.
SQ SEQUENCE 22 AA; 2427 MW; 1D440B2200D4367C CRC64;

Query Match , 11.3%; Score 37; DB 1; Length 22;

Best Local Similarity 58.3%; Pred. No. 2.8e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 36 LWKQDKVAEL 47
||:|:|:|:
DB 2 LWEXIKKANEL 13

RESULT 9
PSBT_MAIZE
ID PSBT_MAIZE STANDARD; PRT; 33 AA.
AC P37257;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Photosystem II reaction center T protein.
GN PSBT OR YCF8.
OS Zea mays (Maize), and
OS Lotus japonicus.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC Clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577, 34305;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Maize;
RX MEDLINE=88210525; PubMed=2835175;
RA Rock C.D., Barkan A., Taylor W.C.;
RT "The maize plastid psbB-psbF-petD gene cluster: spliced and
RT unspliced petB and petD RNAs encode alternative products.";
RL Curr. Genet. 12:69-77(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Maize;
RX MEDLINE=95395841; PubMed=7666415;
RA Maier R.M., Neckermann K., Igloi G.L., Koessel H.;
RT "Complete sequence of the maize chloroplast genome: gene content,
RT hotspots of divergence and fine tuning of genetic information by
RT transcript editing.";
RL J. Mol. Biol. 251:614-628(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=L. japonicus; STRAIN=Accession MG-20;
RX MEDLINE=21082929; PubMed=11214967;
RA Kato T., Kaneko T., Sato S., Nakamura Y., Tabata S.;
RT "Complete structure of the chloroplast genome of a legume, Lotus
RT japonicus.";
RL DNA Res. 7:323-330(2000).
CC -!- FUNCTION: EXACT FUNCTION NOT KNOWN. IT IS ESSENTIAL TO MAINTAIN
CC PHOTOSYNTHETIC ACTIVITY UNDER ADVERSE GROWTH CONDITIONS
CC (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
CC -!- SIMILARITY: BELONGS TO THE PSBT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X05422; -; NOT_ANNOTATED_CDS.
DR EMBL; X85653; CAA60312.1; -.
DR EMBL; AP002983; BAB33223.1; -.
DR MaizeDB; 85519; -.
DR InterPro; IPR001743; PSII_Psbt.
DR Pfam; PF01405; Psbt; 1.
DR ProDom; PD004453; PSII_Psbt; 1.
KW Photosynthesis; Photosystem II; Chloroplast; Transmembrane;
KW Thylakoid.
FT TRANSMEM 3 23 POTENTIAL.

```

SQ  SEQUENCE 33 AA; 3818 MW; A51F72A372756A79 CRC64;
Query Match 11.3%; Score 37; DB 1; Length 33;
Best Local Similarity 31.6%; Pred. No. 4.2e+02;
Matches 6; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 16 GVYGVNVTFFVPSNVPLKE 34
DB 15 GLIIFFAIFFREPKVPTKK 33

RESULT 10
PSBT_TOBAC STANDARD; PRT; 34 AA.
AC P12184;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Photosystem II reaction center T protein.
GN PSBT OR YCF8.
OS Nicotiana tabacum (Common tobacco).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Bright yellow 4;
RA Shinozaki K., Ohme M., Tanaka M., Wakasugi T., Hayashida N.,
RA Matsubayashi T., Zaita N., Chunwongse J., Obokata J.,
RA Yamaguchi-Shinozaki K., Ohto C., Torazawa K., Meng B.-Y., Sugita M.,
RA Dena H., Kamogashira T., Yamada K., Kusuda J., Takaiwa F., Kato A.,
RA Tohoh N., Shimada H., Sugura M.;
RT "The complete nucleotide sequence of the tobacco chloroplast genome:
RT its gene organization and expression.";
RL EMBO J. 5:2043-2049(1986).
CC -!- FUNCTION: EXACT FUNCTION NOT KNOWN. IT IS ESSENTIAL TO MAINTAIN
CC (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
CC -!- SIMILARITY: BELONGS TO THE PSBT FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z00044; CAA77423.2; -
CC InterPro; IPR001743; PSII_Psbt.
CC Pfam; PF01405; Psbt; 1.
CC ProDom; PD004453; PSII_Psbt; 1.
KW Photosynthesis; Photosystem II; Chloroplast; Transmembrane;
KW Thylakoid.
FT TRANSMEM 3 23 POTENTIAL.
SQ SEQUENCE 34 AA; 3932 MW; 29B51F72A372756A CRC64;
Query Match 11.3%; Score 37; DB 1; Length 34;
Best Local Similarity 31.6%; Pred. No. 4.3e+02;
Matches 6; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 16 GVYGVNVTFFVPSNVPLKE 34
DB 15 GLIIFFAIFFREPKVPTKK 33

RESULT 11
YS02_BORBU STANDARD; PRT; 49 AA.
ID YS02_BORBU
AC O54556;

```

```

DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein BBD02/BBH03.
GN BBD02 AND BBH03.
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid lp17 (linear 17 kb) (lp16), and Plasmid lp28-3.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RC MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi.";
RL Nature 390:580-586(1997).
CC -!- SIMILARITY: STRONG, TO B.BURGDORFERI BBF04.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000784; AAC65997.1; -
CC EMBL; AE000793; AAC66357.1; -
CC TIGR; BBD02; -
CC TIGR; BBH03; -
SQ SEQUENCE 49 AA; 5851 MW; E226B54320953273 CRC64;
Query Match 11.3%; Score 37; DB 1; Length 49;
Best Local Similarity 30.0%; Pred. No. 6.2e+02;
Matches 6; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 38 KKQDKVAELENSEFRAFSS 57
DB 3 KKQKNKSEIEKTELEIINN 22

RESULT 12
YU04_BORBU STANDARD; PRT; 49 AA.
ID YU04_BORBU
AC O51014;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein BBF04.
GN BBF04.
OS Borrelia burgdorferi (Lyme disease spirochete).
OG Plasmid lp28-1.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RC MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,

```

```
RA Garland S., Fujil C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi.",
RL Nature 390:580-586(1997).
CC -1- SIMILARITY: STRONG. TO B.BURGDORFERI BBD02 AND BBH03.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE000794; AAC66385.1; -.
CC TIGR: BBF04; -.
CC KW Hypothetical protein; Plasmid; Complete proteome.
CC SQ SEQUENCE 49 AA; 5808 MW; E226B544F12F3273 CRC64;
Query Match 11.3%; Score 37; DB 1; Length 49;
Best Local Similarity 30.0%; Pred. No. 6.2e+02;
Matches 6; Conservative 7; Mismatches 7; Indels 0; Gaps 0;
Qy 38 KKOKDKVALENSERAPSS 57
Db 3 KNQKNKCSIEKTOLEIINT 22
RESULT 13
SYD_MACFA STANDARD; PRT; 55 AA.
AC P36391.
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Sex-determining region Y protein ('Testis-determining factor')
DE (Fragment).
GN SRY OR TDF.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94214723; PubMed=8162230;
RA Griffiths R., Tiwari B.;
RT "Primers for the differential amplification of the sex-determining
RT region Y gene in a range of mammal species.";
RL Mol. Ecol. 2:405-406(1993).
CC -1- FUNCTION: TRANSCRIPTIONAL ACTIVATOR WHICH REGULATES A GENETIC
CC SWITCH IN MALE DEVELOPMENT. IT IS RESPONSIBLE FOR INITIATING MALE
CC SEX DETERMINATION. SRY HMG BOX RECOGNIZES DNA BY PARTIAL
CC INTERCALATION IN THE MINOR GROOVE.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: CONTAINS 1 HMG BOX.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z26906; CAA81532.1; -.
CC HSP: Q05066; 1HRV.
CC InterPro: IPR000910; HMG_box.1.
CC Pfam: PF00505; HMG_box.1.
CC SMART: SM00398; HMG; 1.
CC KW DNA-binding; Nuclear protein; Transcription regulation; Activator;
CC Sexual differentiation.
```

```
FT NON_TER 1 1 HMG_BOX.
FT DNA_BIND <1 >55
FT NON_TER 55
SQ SEQUENCE 55 AA; 6839 MW; 34A3CB029AC93233 CRC64;
Query Match 11.3%; Score 37; DB 1; Length 55;
Best Local Similarity 26.3%; Pred. No. 7e+02;
Matches 5; Conservative 7; Mismatches 7; Indels 0; Gaps 0;
Qy 35 VLWKKQKDKVALENSER 53
Db 1 IVWSRQRRKMDLENPKMR 19
RESULT 14
TAT_HV1B5 STANDARD; PRT; 58 AA.
AC P04612;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE TAT protein (Transactivating regulatory protein) (Fragment).
DE TAT.
GN TAT.
OS Human immunodeficiency virus type 1 (HIV isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11682;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85111123; PubMed=2578615;
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
RA Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,
RA Baesecker K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,
RA Lautenberger J.A., Papas T.S., Grayeb J., Chang N.T., Gallo R.C.,
RA Wong-Staal F.;
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
RT Nature 313:277-284(1985).
RL -1- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC TRANS-ACTIVATING RESPONSE SEQUENCE (TAR) RNA ELEMENT AND
CC ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC PROMOTER.
CC -1- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear; nucleolar.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: K02012; AAA44656.1; -.
CC HIV: K02012; TATSBH5.
CC InterPro: IPR001831; HIV_Tat.
CC Pfam: PF00539; Tat; 1.
CC KW Transcription regulation; Activator; RNA-binding; Nuclear protein;
CC AIDS.
FT NON_TER 58 58
SQ SEQUENCE 58 AA; 6800 MW; E36C21F8FFD813E3 CRC64;
Query Match 11.3%; Score 37; DB 1; Length 58;
Best Local Similarity 38.9%; Pred. No. 7.4e+02;
Matches 7; Conservative 2; Mismatches 9; Indels 0; Gaps 0;
Qy 3 CFGFISCFSQOIVGVVYG 20
Db 31 CFHCQVCFTKALGISYG 48
RESULT 15
SODC_STRHE STANDARD; PRT; 31 AA.
ID SODC_STRHE
AC F81163;
```

DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1) (Fragments).
 OS Striga hermonthica.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Lamiales; Orobanchaceae; Striga.
 OX NCBI_TaxID=68872;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Seedling;
 RA Stranger A., Corbett J.M., Dunn M.J., Totty N.F., Sterling A.,
 RA Bolwell G.P.;
 RT "Identification of developmentally-specific markers in germinating and
 RT haustorial stages of Striga hermonthica (Del.) Benth. seedlings.";
 RL J. Exp. Bot. 50:269-274(1999).
 CC -!- FUNCTION: Destroys radicals which are normally produced within the
 CC cells and which are toxic to biological systems.
 CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
 CC -!- COFACTOR: Copper and zinc (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- DEVELOPMENTAL STAGE: PREFERENTIALLY EXPRESSED IN GERMINATING
 CC SEEDLINGS.
 CC -!- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.
 DR InterPro: IPR001424; SOD_CU_ZN.
 DR PROSITE: PS00087; SOD_CU_ZN_1; PARTIAL.
 DR PROSITE: PS00332; SOD_CU_ZN_2; PARTIAL.
 KW Oxidoreductase; Copper; Zinc.
 FT NON_TER 1 1
 FT NON_TER 17 18
 FT NON_TER 31 31
 SQ SEQUENCE 31 AA; 3109 MW; 40768DE8F8FC2958 CRC64;

Query Match 11.0%; Score 36; DB 1; Length 31;
 Best local Similarity 46.7%; Pred. No. 5.2e+02;
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 18 YVGNVTFHVPFSNVPL 32
 | | :|| | :||
 Db 11 VAGTITFIADSQIPL 25

Search completed: January 28, 2003, 08:59:43
 Job time : 12 secs

11

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: January 28, 2003, 08:58:05 ; Search time 28 seconds
(without alignments)
448.888 Million cell updates/sec

Title: US-09-730-465-2_COPY_20_80
Sequence: 1 LHCFFGSCFSQIYGVYGV.....DKVAELENSEFRAFSFKNR 61

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 46985

Minimum DB seq length: 0
Maximum DB seq length: 61

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	47	14.4	19	11 Q9QV70	Q9qV70 rattus sp.
2	45	13.8	60	7 Q98239	Q98239 tapirus bai
3	45	13.8	60	7 Q98240	Q98240 tapirus ind
4	45	13.8	60	7 Q98241	Q98241 diceros bic
5	45	13.8	60	7 Q98242	Q98242 diceros bic
6	45	13.8	60	7 Q98243	Q98243 diceros bic
7	45	13.8	60	7 Q98244	Q98244 ceratotheri
8	45	13.8	60	7 Q98246	Q98246 rhinoceros
9	44.5	13.6	42	10 Q08574	Q08574 catharanthu
10	43	13.1	53	6 P82126	P82126 sus scrofa
11	43	13.1	55	5 Q9G292	Q9g292 plasmodium
12	43	13.1	56	5 Q9UA59	Q9ua59 caenorhabdi
13	43	13.1	58	5 Q9G298	Q9g298 plasmodium
14	42.5	13.0	48	9 Q9AJB5	Q9ajb5 streptococc
15	42	12.8	40	2 Q45682	Q45682 bacillus su
16	42	12.8	47	7 Q95600	Q95600 squalus aca

17	42	12.8	54	10 Q40809	Q40809 physalis cr
18	42	12.8	55	2 Q9F2C3	Q9f2c3 salmonella
19	42	12.8	59	2 Q03122	Q03122 streptococc
20	42	12.8	60	7 Q98245	Q98245 ceratotheri
21	42	12.8	60	16 Q9JQU1	Q9jqu1 neisseria m
22	41.5	12.7	35	10 Q9S9G6	Q9s9g6 cucurbita (
23	41.5	12.7	43	16 Q9K248	Q9k248 chlamydia p
24	41.5	12.7	57	16 Q9PPD3	Q9ppd3 campylobact
25	41	12.5	47	4 Q9UN35	Q9un35 homo sapien
26	41	12.5	47	8 Q9SCN8	Q9scn8 latrobea hi
27	41	12.5	60	16 Q9RW69	Q9rw69 deinococcus
28	40.5	12.4	54	4 Q96LF3	Q96lf3 homo sapien
29	40	12.2	46	8 Q95CT7	Q95ct7 gastrolobiu
30	40	12.2	46	8 Q95CT5	Q95ct5 gastrolobiu
31	40	12.2	46	8 Q95CT2	Q95ct2 gastrolobiu
32	40	12.2	46	8 Q95CS2	Q95cs2 gastrolobiu
33	40	12.2	46	8 Q95CF5	Q95cf5 gastrolobiu
34	40	12.2	46	8 Q95C09	Q95c09 gastrolobiu
35	40	12.2	46	8 Q95CQ7	Q95cq7 gastrolobiu
36	40	12.2	46	8 Q95CQ4	Q95cq4 gastrolobiu
37	40	12.2	46	8 Q95CQ0	Q95cq0 gastrolobiu
38	40	12.2	46	8 Q95CP9	Q95cp9 gastrolobiu
39	40	12.2	46	8 Q95CP8	Q95cp8 gastrolobiu
40	40	12.2	47	8 Q95CU7	Q95cu7 brachysema
41	40	12.2	47	8 Q95CU6	Q95cu6 brachysema
42	40	12.2	47	8 Q95CU5	Q95cu5 brachysema
43	40	12.2	47	8 Q95CU4	Q95cu4 brachysema
44	40	12.2	47	8 Q95CU3	Q95cu3 brachysema
45	40	12.2	47	8 Q95CU1	Q95cu1 brachysema

ALIGNMENTS

RESULT 1

Q9QV70 PRELIMINARY: PRT: 19 AA.
ID Q9QV70
AC Q9QV70;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE Lactase-phlorizin hydrolase (Fragment).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE.
RX MEDLINE=93293888; PubMed=8514793;
RA Dudley M.A., Hachey D.L., Quaroni A., Hutchens T.W., Nichols B.L.,
RA Rosenberger J., Perkinson J.S., Cook G., Reeds P.J.;
RT "In vivo sucrose-isomaltase and lactase-phlorizin hydrolase turnover
RT in the fed adult rat";
RL J. Biol. Chem. 268:13609-13616(1993);
SQ SEQUENCE 19 AA; 2174 MW; 136CC0A8155A1D07 CRC64;

Query Match 14.4%; Score 47; DB 11; Length 19;
Best Local Similarity 44.4%; Pred. No. 48;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 22 VTFHVPSPVPLKEVLNKK 39

DB 1 VTDSLPEVPKAKIVWEK 18

RESULT 2

Q98239 PRELIMINARY: PRT: 60 AA.
ID Q98239
AC Q98239;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MHC class II DR-alpha (Fragment).

```
GN DRA.
OS Tapirus bairdii (Baird's tapir).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Tapiridae; Tapirus.
OX NCBI_TaxID=56117;
RN [1]
RP SEQUENCE FROM N.A.
RA Fraser D.G., Ryder O.A.;
RT "DRA polymorphism in the Perissodactyla.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF113549; AAD19966.1; -.
DR InterPro; IPR001003; MHC_II_alpha.
DR Pfam; PF00993; MHC_II_alpha; 1.
FT NON_TER 1
FT NON_TER 60
SQ SEQUENCE 60 AA; 6872 MW; 192850C89D3F4A91 CRC64;

Query Match 13.8%; Score 45; DB 7; Length 60;
Best Local Similarity 30.0%; Pred. No. 3.1e+02;
Matches 12; Conservative 8; Mismatches 8; Indels 12; Gaps 2;

Qy 20 GNVTFFHVPNSVPLKEVLWKQKDKVAELENSEFRASFSEFK 59
Db 14 GDEIFHV--DMEKKEVTWRLE-----EFGRFASFE 41

RESULT 3
O98240 ID O98240 PRELIMINARY; PRT; 60 AA.
AC O98240;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MHC class II DR-alpha (Fragment).
GN DRA.
OS Tapirus indicus (Asiatic tapir) (Malayan tapir).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Tapiridae; Tapirus.
OX NCBI_TaxID=9802;
RN [1]
RP SEQUENCE FROM N.A.
RA Fraser D.G., Ryder O.A.;
RT "DRA polymorphism in the Perissodactyla.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF113549; AAD19967.1; -.
DR InterPro; IPR001003; MHC_II_alpha.
DR Pfam; PF00993; MHC_II_alpha; 1.
FT NON_TER 1
FT NON_TER 60
SQ SEQUENCE 60 AA; 6856 MW; 053450D4813F4A91 CRC64;

Query Match 13.8%; Score 45; DB 7; Length 60;
Best Local Similarity 30.0%; Pred. No. 3.1e+02;
Matches 12; Conservative 8; Mismatches 8; Indels 12; Gaps 2;

Qy 20 GNVTFFHVPNSVPLKEVLWKQKDKVAELENSEFRASFSEFK 59
Db 14 GDEIFHV--DMEKKEVTWRLE-----EFGRFASFE 41

RESULT 4
O98241 ID O98241 PRELIMINARY; PRT; 60 AA.
AC O98241;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MHC class II DR-alpha (Fragment).
GN DRA.
OS Diceros bicornis (Black rhinoceros).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Diceros.
OX NCBI_TaxID=9805;
RN [1]
RP SEQUENCE FROM N.A.
RA Fraser D.G., Ryder O.A.;
RT "DRA polymorphism in the Perissodactyla.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
```

```
RN [1]
RP SEQUENCE FROM N.A.
RA Fraser D.G., Ryder O.A.;
RT "DRA polymorphism in the Perissodactyla.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF113549; AAD19968.1; -.
DR InterPro; IPR001003; MHC_II_alpha.
DR Pfam; PF00993; MHC_II_alpha; 1.
FT NON_TER 1
FT NON_TER 60
SQ SEQUENCE 60 AA; 6870 MW; 053450D5313F4A91 CRC64;

Query Match 13.8%; Score 45; DB 7; Length 60;
Best Local Similarity 30.0%; Pred. No. 3.1e+02;
Matches 12; Conservative 8; Mismatches 8; Indels 12; Gaps 2;

Qy 20 GNVTFFHVPNSVPLKEVLWKQKDKVAELENSEFRASFSEFK 59
Db 14 GDEIFHV--DMEKKEVTWRLE-----EFGRFASFE 41

RESULT 5
O98242 ID O98242 PRELIMINARY; PRT; 60 AA.
AC O98242;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MHC class II DR-alpha (Fragment).
GN DRA.
OS Diceros bicornis (Black rhinoceros).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Diceros.
OX NCBI_TaxID=9805;
RN [1]
RP SEQUENCE FROM N.A.
RA Fraser D.G., Ryder O.A.;
RT "DRA polymorphism in the Perissodactyla.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF113550; AAD19969.1; -.
DR InterPro; IPR001003; MHC_II_alpha.
DR Pfam; PF00993; MHC_II_alpha; 1.
FT NON_TER 1
FT NON_TER 60
SQ SEQUENCE 60 AA; 6870 MW; 053450D5313F4A91 CRC64;

Query Match 13.8%; Score 45; DB 7; Length 60;
Best Local Similarity 30.0%; Pred. No. 3.1e+02;
Matches 12; Conservative 8; Mismatches 8; Indels 12; Gaps 2;

Qy 20 GNVTFFHVPNSVPLKEVLWKQKDKVAELENSEFRASFSEFK 59
Db 14 GDEIFHV--DMEKKEVTWRLE-----EFGRFASFE 41

RESULT 6
O98243 ID O98243 PRELIMINARY; PRT; 60 AA.
AC O98243;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MHC class II DR-alpha (Fragment).
GN DRA.
OS Diceros bicornis (Black rhinoceros).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Diceros.
OX NCBI_TaxID=9805;
RN [1]
RP SEQUENCE FROM N.A.
RA Fraser D.G., Ryder O.A.;
RT "DRA polymorphism in the Perissodactyla.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL: AF113551; AAD19970.1; -
DR InterPro; IPR001003; MHC_II_alpha.
DR Pfam; PF00993; MHC_II_alpha; 1.
FT NON_TER 1
FT NON_TER 60
SQ SEQUENCE 60 AA; 6870 MW; 053450D5313F4A91 CRC64;

Query Match
Best Local Similarity 13.8%; Score 45; DB 7; Length 60;
Matches 12; Conservative 8; Mismatches 8; Indels 12; Gaps 2;

QY 20 GNVTFHVPNSVPLKEVLWKQKDKVAELENSEFRFASFSFK 59
Db 14 GDEIFHV--DMEKKEVTWRLE-----EFGRFASFE 41

RESULT 7
O98244 PRELIMINARY; PRT; 60 AA.
AC O98244;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MHC class II DR-alpha (Fragment).
GN DRA.
OS Ceratotherium sinu (White rhinoceros) (Square-lipped rhinoceros).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Ceratotherium.
OX NCBI_TaxID=9807;
RN [1]
RP SEQUENCE FROM N.A.
RA Fraser D.G., Ryder O.A.;
RT "DRA polymorphism in the Perissodactyla.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF113552; AAD19971.1; -
DR InterPro; IPR001003; MHC_II_alpha.
DR Pfam; PF00993; MHC_II_alpha; 1.
FT NON_TER 1
FT NON_TER 60
SQ SEQUENCE 60 AA; 6870 MW; 053450D5313F4A91 CRC64;

Query Match
Best Local Similarity 13.8%; Score 45; DB 7; Length 60;
Matches 12; Conservative 8; Mismatches 8; Indels 12; Gaps 2;

QY 20 GNVTFHVPNSVPLKEVLWKQKDKVAELENSEFRFASFSFK 59
Db 14 GDEIFHV--DMEKKEVTWRLE-----EFGRFASFE 41

RESULT 8
O98246 PRELIMINARY; PRT; 60 AA.
AC O98246;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MHC class II DR-alpha (Fragment).
GN DRA.
OS Rhinoceros unicornis (Greater Indian rhinoceros).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Rhinoceros.
OX NCBI_TaxID=9809;
RN [1]
RP SEQUENCE FROM N.A.
RA Fraser D.G., Ryder O.A.;
RT "DRA polymorphism in the Perissodactyla.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF113554; AAD19973.1; -
DR InterPro; IPR001003; MHC_II_alpha.
DR Pfam; PF00993; MHC_II_alpha; 1.
FT NON_TER 1
FT NON_TER 60
```

```
SQ SEQUENCE 60 AA; 6868 MW; BF2150D52634269B CRC64;

Query Match
Best Local Similarity 13.8%; Score 45; DB 7; Length 60;
Matches 12; Conservative 8; Mismatches 8; Indels 12; Gaps 2;

QY 20 GNVTFHVPNSVPLKEVLWKQKDKVAELENSEFRFASFSFK 59
Db 14 GDEIFHV--DLEKKEVTWRLE-----EFGRFASFE 41

RESULT 9
Q08574 PRELIMINARY; PRT; 42 AA.
AC Q08574;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cytochrome P-450 (EC 1.14.14.1) (CRL3) (Fragment).
GN P450CRL3.
OS Catharanthus roseus (Rosy periwinkle) (Madagascar periwinkle).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Gentianales; Apocynaceae; Rauvolfioideae;
OC Vinceae; Catharanthus.
OX NCBI_TaxID=4058;
RP SEQUENCE FROM N.A.
RC STRAIN=G.DON;
RX MEDLINE=93283641; PubMed=8507838;
RA Meijer A.H., Souer E., Verpoorte R., Hoge J.H.C.;
RT "Isolation of cytochrome P-450 cDNA clones from the higher plant
Catharanthus roseus by a PCR strategy.";
RL Plant Mol. Biol. 22:379-383(1993).
CC -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
MONOOXYGENASES. MAY BE A GERANIOL-10-HYDROXYLASE. -!- ROH +
-!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +
OXIDIZED FLAVOPROTEIN + H(2)O.
CC -!- PATHWAY: INDOLE ALKALOID SYNTHESIS.
CC -!- SUBCELLULAR LOCATION: MEMBRANE BOUND.
CC -!- SIMILARITY: MEMBER OF THE CYTOCHROME P-450 FAMILY.
DR EMBL: X65785; CAA49440.1; -
KW Oxidoreductase; Monooxygenase; Electron transport; Transmembrane;
KW Heme.
FT NON_TER 1
FT NON_TER 42
SQ SEQUENCE 42 AA; 4641 MW; 23C83C9ABA4B34B4 CRC64;

Query Match
Best Local Similarity 13.6%; Score 44.5; DB 10; Length 42;
Matches 11; Conservative 11; Mismatches 10; Indels 5; Gaps 2;

QY 22 VTFHVPNSVPLKEVL---WKKQKDKVAELENSEFR 53
Db 5 MSFAIPNVTPLAQLLLHFDWKSAYGKLEDLDMTGAR 41

RESULT 10
P82126 PRELIMINARY; PRT; 53 AA.
AC P82126;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Neuronal protein NP-190 (Fragments).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9825;
RN [1]
RP SEQUENCE, AND CHARACTERIZATION.
RC TISSUE=FETAL BRAIN;
RX PubMed=9538211;
RA Ho S.-C., Wakatsuki S., Arioka M., Yamasaki M., Kitamoto K.;
```


OS Streptococcus thermophilus bacteriophage DT1.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=90410;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=DT1;

RX MEDLINE=99160757; PubMed=10049822;

RA Tremblay D.M., Moineau S.;

RT "Complete genomic sequence of the lytic bacteriophage DT1 of

Streptococcus thermophilus.";

RL Virology 255:63-76(1999).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=DT1;

RX MEDLINE=21382762; PubMed=11489121;

RA Duplessis M., Moineau S.;

RT "Identification of a genetic determinant responsible for host

specificity in Streptococcus thermophilus bacteriophages.";

RL Mol. Microbiol. 41:325-336(2001).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=DT1;

RA Tremblay D.M., Moineau S.;

RL Submitted (AUG-1998) to the EMBL/GenBank/DBDJ databases.

DR EMBL; AF085222; AAD21898.1; --

KW Hypothetical protein.

SQ SEQUENCE 48 AA; 5579 MW; 549B8C06D27967B2 CRC64;

Query Match 13.0%; Score 42.5; DB 9; Length 48;

Best Local Similarity 37.1%; Pred. No. 4.9e+02;

Matches 13; Conservative 6; Mismatches 11; Indels 5; Gaps 2;

OY 17 VYGVNTHVPSNVPLKEVLWKQKDKVAELENSE 51

DB 10 IVNNYTF--KRVF--KVLKPKVKELIADMYNDE 39

RESULT 15

Q45682

ID Q45682 PRELIMINARY; PRT; 40 AA.

AC Q45682;

DT 01-NOV-1996 (Tremblrel. 01, Created)

DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

DE TagA, tagB, tagC and tagD genes.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;

OC Bacillaceae; Bacillus.

OX NCBI_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=H168;

RX MEDLINE=91311391; PubMed=1906926;

RA Mael C., Young M.W., Karamata D.;

RT "Genes concerned with biosynthesis of polyglycerol phosphate in

Bacillus subtilis are organised in two divergently transcribed

operons.";

RL J. Gen. Microbiol. 137:929-941(1991).

DR EMBL; M57497; AAA22846.1; --

SQ SEQUENCE 40 AA; 4769 MW; 03000EF83FE3FCC0 CRC64;

Query Match 12.8%; Score 42; DB 2; Length 40;

Best Local Similarity 66.7%; Pred. No. 4.6e+02;

Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 3 CFGFISCFSQOI 14

DB 22 CFNVISCLWQOI 33

Search completed: January 28, 2003, 09:00:17

Job time : 29 secs



GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 08:55:00 ; Search time 35 Seconds
(without alignments)
232.237 Million cell updates/sec

Title: us-09-730-465-2_COPY_20_80

Perfect score: 327

Sequence: 1 LUCFGIFCSQQIYGVVYG.....DKVAELENSEFRASFKNR 61

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 477368

Minimum DB seq length: 0

Maximum DB seq length: 61

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	28.1	38	22	AA84103
2	92	28.1	38	22	AA84104
3	90	27.5	35	22	AA84106
4	87	26.6	17	22	AA84077
5	87	26.6	32	22	AA84085
6	87	26.6	36	22	AA84102
7	87	26.6	42	22	AA84105
8	82	25.1	42	22	AA84101
9	76	23.2	29	9	AA82988
10	73	22.3	29	9	AA82986

11	71	21.7	29	9	AA82989	Sheep erythrocyte
12	64	19.6	29	9	AA82987	Sheep erythrocyte
13	62	19.0	29	9	AA82986	Sheep erythrocyte
14	62	19.0	29	15	AA860315	SEQ ID NO 454 from
15	55	16.8	57	20	AA19736	Human polypeptide
16	51	15.6	56	22	AA08362	Human peptide #415
17	50.5	15.4	56	22	AB22764	Human peptide #415
18	50.5	15.4	56	22	AB228193	Human peptide #441
19	50.5	15.4	56	22	AB332935	Peptide #441 encod
20	50.5	15.4	56	22	AB33368	Peptide #474 encod
21	50.5	15.4	56	22	AB18414	Protein #413 encod
22	50.5	15.4	56	22	AB18827	Human brain expres
23	50.5	15.4	56	22	AA53736	Human brain expres
24	50.5	15.4	56	22	AA54153	Human bone marrow
25	50.5	15.4	56	22	AA66119	Human bone marrow
26	50.5	15.4	56	22	AA66547	Peptide #424 encod
27	50.5	15.4	56	22	AA13990	Peptide #854 encod
28	50.5	15.4	56	22	AA14420	Peptide #433 encod
29	50.5	15.4	56	22	AA26396	Peptide #870 encod
30	50.5	15.4	56	22	AA26833	Peptide #413 encod
31	50.5	15.4	56	22	AA01731	Peptide #829 encod
32	50.5	15.4	56	22	AA02147	Human peptide enco
33	50.5	15.4	56	23	ABG35769	Human peptide enco
34	50.5	15.4	56	23	ABG36199	Human OREX protein
35	50.5	15.4	57	23	ABP07351	Peptide effecting
36	48	14.7	29	19	AA39971	Human cardiovascul
37	48	14.7	57	22	AA22461	BBC1 derived, pres
38	46	14.1	36	16	AA87199	G-protein coupled
39	46	14.1	44	15	AA50764	G-protein coupled
40	46	14.1	44	15	AA50765	G-protein coupled
41	46	14.1	44	17	AA02956	G-protein coupled
42	46	14.1	44	17	AA02957	G-protein coupled
43	46	14.1	45	15	AA50766	G-protein coupled
44	46	14.1	45	17	AA02958	G-protein coupled
45	46	14.1	47	15	AA50767	G-protein coupled

ALIGNMENTS

RESULT 1

AA84103

ID AA84103 standard; peptide; 38 AA.

XX AA84103;

XX 06-AUG-2001 (first entry)

XX Immunomodulatory peptide used to treat allergic responses.

XX Immunomodulatory peptide; immune response; autoimmune disease; allergy; asthma; host-versus-graft rejection; T cell; anergy; apoptosis.

XX Synthetic.

OS Cryptomeria japonica.

XX Key

XX Location/Qualifiers

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

```
XX
PI Zimmerman DH;
XX
DR WPI: 2001-374498/39.
XX
PT Novel immunomodulatory peptide construct useful for modulating an
PT inappropriate immune response in an individual at risk for autoimmune
PT disease, allergic reactions, asthma or host-graft or graft-host disease
PT
XX
PS Claim 4; Page 26; 55pp; English.
XX
CC The specification describes an immunomodulatory peptide construct.
CC The immunomodulatory peptide comprises a first peptide associated with
CC autoimmune disease, allergy or asthma, or host-versus-graft rejection
CC and which will bind to an antigen receptor on a set or subset of T cells,
CC linked to a second immune modulating peptide which will cause a directed
CC immune response by the set or subset of T cells to which the first
CC peptide is attached. Alternatively, the second peptide will bind to a
CC T cell receptor site on the surface of the T cell which will cause the
CC set or subset of T cells to initiate, but not complete, an immune
CC response to cause the set or subset of T cells to undergo anergy and
CC apoptosis. The immunomodulatory peptides are useful for eliminating a set
CC or subset of T cells involved in autoimmune response. They are useful for
CC the treatment of autoimmune disease, allergic reactions, asthma or
CC host-graft or graft-host rejection. The immunomodulatory peptides are
CC also useful for interrupting an autoimmune disease associated pathway
CC necessary to complete T cell activation. The present sequence represents
CC an immunomodulatory peptide used to treat allergic responses.
XX
XX
SQ Sequence 38 AA;
XX
Query Match 28.1%; Score 92; DB 22; Length 38;
Best Local Similarity 80.0%; Pred. No. 0.00013;
Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 35 VLWKKQKDKVAELENSEFRFSSPK 59
DB 1 VLWKKQKDKVAELENSEGGGSSMK 25
RESULT 2
AAB84104
ID AAB84104 standard; peptide; 38 AA.
XX
AC AAB84104;
XX
DT 06-AUG-2001 (first entry)
XX
DE Immunomodulatory peptide used to treat allergic responses.
XX
KW Immunomodulatory peptide; immune response; autoimmune disease; allergy;
KW asthma; host-versus-graft rejection; T cell; anergy; apoptosis.
XX
OS Synthetic.
OS Cryptomeria japonica.
XX
FH Key Location/Qualifiers
FT Peptide 1..17
FT Peptide /note= "LFA-3 peptide"
FT Peptide 18..23
FT Peptide /note= "linker"
FT Peptide 24..38
FT Peptide /note= "pollen antigen"
XX
PN WO200136448-A2.
XX
PD 25-MAY-2001.
XX
PF 27-OCT-2000; 2000WO-US41646.
XX
PR 27-OCT-1999; 99US-0161734.
XX
(CELS-) CEL-SCI CORP.
Zimmerman DH;
WPI: 2001-374498/39.
Novel immunomodulatory peptide construct useful for modulating an
inappropriate immune response in an individual at risk for autoimmune
disease, allergic reactions, asthma or host-graft or graft-host disease
Claim 4; Page 26; 55pp; English.
The specification describes an immunomodulatory peptide construct.
The immunomodulatory peptide comprises a first peptide associated with
autoimmune disease, allergy or asthma, or host-versus-graft rejection
and which will bind to an antigen receptor on a set or subset of T cells,
linked to a second immune modulating peptide which will cause a directed
immune response by the set or subset of T cells to which the first
peptide is attached. Alternatively, the second peptide will bind to a
T cell receptor site on the surface of the T cell which will cause the
set or subset of T cells to initiate, but not complete, an immune
response to cause the set or subset of T cells to undergo anergy and
apoptosis. The immunomodulatory peptides are useful for eliminating a set
or subset of T cells involved in autoimmune response. They are useful for
the treatment of autoimmune disease, allergic reactions, asthma or
host-graft or graft-host rejection. The immunomodulatory peptides are
also useful for interrupting an autoimmune disease associated pathway
necessary to complete T cell activation. The present sequence represents
an immunomodulatory peptide used to treat allergic responses.
Query Match 28.1%; Score 92; DB 22; Length 38;
Best Local Similarity 74.1%; Pred. No. 0.00013;
Matches 20; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
QY 35 VLWKKQKDKVAELENSEFRFSSFKNR 61
DB 1 VLWKKQKDKVAELENSEGGGSSIASR 27
RESULT 3
AAB84106
ID AAB84106 standard; peptide; 35 AA.
XX
AC AAB84106;
XX
DT 06-AUG-2001 (first entry)
XX
DE Immunomodulatory peptide used to treat allergic responses.
XX
KW Immunomodulatory peptide; immune response; autoimmune disease; allergy;
KW asthma; host-versus-graft rejection; T cell; anergy; apoptosis.
XX
OS Synthetic.
OS Apis sp.
XX
FH Key Location/Qualifiers
FT Peptide 1..17
FT Peptide /note= "LFA-3 peptide"
FT Peptide 18..23
FT Peptide /note= "linker"
FT Peptide 24..35
FT Peptide /note= "bee venom antigen"
XX
PN WO200136448-A2.
XX
PD 25-MAY-2001.
XX
PF 27-OCT-2000; 2000WO-US41646.
XX
PR 27-OCT-2000; 2000WO-US41646.
```



```

PF 27-OCT-2000; 2000WO-US41646.
XX
PR 27-OCT-1999; 99US-0161734.
XX
PA (CELS-) CEL-SCI CORP.
XX
PI Zimmerman DH;
XX
DR WPI; 2001-374498/39.
XX
PT Novel immunomodulatory peptide construct useful for modulating an
PT inappropriate immune response in an individual at risk for autoimmune
PT disease, allergic reactions, asthma or host-graft or graft-host disease
PT
PS Claim 4; Page 22; 55pp; English.
XX
CC The specification describes an immunomodulatory peptide construct.
CC The immunomodulatory peptide comprises a first peptide associated with
CC autoimmune disease, allergy or asthma, or host-versus-graft rejection
CC and which will bind to an antigen receptor on a set or subset of T cells,
CC linked to a second immune modulating peptide which will cause a directed
CC immune response by the set or subset of T cells to which the first
CC peptide is attached. Alternatively, the second peptide will bind to a
CC T cell receptor site on the surface of the T cell which will cause the
CC set or subset of T cells to initiate, but not complete, an immune
CC response to cause the set or subset of T cells to undergo energy and
CC apoptosis. The immunomodulatory peptides are useful for eliminating a set
CC or subset of T cells involved in autoimmune response. They are useful for
CC the treatment of autoimmune disease, allergic reactions, asthma or
CC host-graft or graft-host rejections. The immunomodulatory peptides are
CC also useful for interrupting an autoimmune disease associated pathway
CC necessary to complete T cell activation. The present peptide represents
CC an immunomodulatory peptide of the invention, which is used to treat
CC Addison's disease. It comprises conformational epitopes from steroid 21
CC hydroxylase fused, via linkers, to the LFA-3 peptide.
XX
SQ Sequence 32 AA;
XX
Query Match 26.6%; Score 87; DB 22; Length 32;
Best Local Similarity 100.0%; Pred. NO. 0.00047;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 35 VLWKKQDKVAELENSE 51
DB 1 VLWKKQDKVAELENSE 17
RESULT 6
AAB84102
ID AAB84102 standard; peptide; 36 AA.
XX
AC AAB84102;
XX
DT 06-AUG-2001 (first entry)
XX
DE Immunomodulatory peptide used to treat allergic responses.
XX
KW Immunomodulatory peptide; immune response; autoimmune disease; allergy;
KW asthma; host-versus-graft rejection; T cell; anergy; apoptosis.
XX
OS Synthetic.
OS Cryptomeria japonica.
XX
FH Key Location/Qualifiers
FT Peptide 1..17
FT Peptide /note= "LFA-3 peptide"
FT Peptide 18..23
FT Peptide /note= "linker"
FT Peptide 24..36
FT Peptide /note= "pollen antigen"
XX
PN WO200136448-A2.
XX
25-MAY-2001.
XX
27-OCT-2000; 2000WO-US41646.
XX
27-OCT-1999; 99US-0161734.
XX
PA (CELS-) CEL-SCI CORP.
XX
PI Zimmerman DH;
XX
DR WPI; 2001-374498/39.
XX
PT Novel immunomodulatory peptide construct useful for modulating an
PT inappropriate immune response in an individual at risk for autoimmune
PT disease, allergic reactions, asthma or host-graft or graft-host disease
PT
PS Claim 4; Page 26; 55pp; English.
XX
CC The specification describes an immunomodulatory peptide construct.
CC The immunomodulatory peptide comprises a first peptide associated with
CC autoimmune disease, allergy or asthma, or host-versus-graft rejection
CC and which will bind to an antigen receptor on a set or subset of T cells,
CC linked to a second immune modulating peptide which will cause a directed
CC immune response by the set or subset of T cells to which the first
CC peptide is attached. Alternatively, the second peptide will bind to a
CC T cell receptor site on the surface of the T cell which will cause the
CC set or subset of T cells to initiate, but not complete, an immune
CC response to cause the set or subset of T cells to undergo energy and
CC apoptosis. The immunomodulatory peptides are useful for eliminating a set
CC or subset of T cells involved in autoimmune response. They are useful for
CC the treatment of autoimmune disease, allergic reactions, asthma or
CC host-graft or graft-host rejections. The immunomodulatory peptides are
CC also useful for interrupting an autoimmune disease associated pathway
CC necessary to complete T cell activation. The present peptide represents
CC an immunomodulatory peptide of the invention, which is used to treat
CC Addison's disease. It comprises conformational epitopes from steroid 21
CC hydroxylase fused, via linkers, to the LFA-3 peptide.
XX
SQ Sequence 35 AA;
XX
Query Match 26.6%; Score 87; DB 22; Length 36;
Best Local Similarity 100.0%; Pred. NO. 0.00054;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 35 VLWKKQDKVAELENSE 51
DB 1 VLWKKQDKVAELENSE 17
RESULT 7
AAB84105
ID AAB84105 standard; peptide; 42 AA.
XX
AC AAB84105;
XX
DT 06-AUG-2001 (first entry)
XX
DE Immunomodulatory peptide used to treat allergic responses.
XX
KW Immunomodulatory peptide; immune response; autoimmune disease; allergy;
KW asthma; host-versus-graft rejection; T cell; anergy; apoptosis.
XX
OS Synthetic.
OS Cryptomeria japonica.
XX
FH Key Location/Qualifiers
FT Peptide 1..17
FT Peptide /note= "LFA-3 peptide"
FT Peptide 18..23
FT Peptide /note= "linker"
FT Peptide 24..42
FT Peptide /note= "pollen antigen"
XX

```



```

PF 19-NOV-1987; 87EP-0008695.
XX
PR 22-NOV-1986; 86DE-3639920.
XX
PA (FARH ) HOECHST AG.
XX
PI Hunig T, Tiefenthaler G, Mitnacht R, Meuer S;
XX
DR WPI; 1988-148827/22.
XX
PT Homogeneous sheep TILTS and human TILTS glycoprotein - useful as
XX immunomodulators and for monoclonal antibody prodn.
XX
PS Claim 1; Page 9; 9pp; German.
XX
CC Specific example of N-terminal sequence of sheep TILTS. The entire
XX TILTS protein is used for immunising an animal and antibodies are
CC isolated from the serum. These antibodies are in turn used to react
CC with human TILTS glycoprotein in an immunoassay. Both the sheep and
CC human derived proteins are useful as immunomodulators.
CC See also AAP80286, AAP82986-P82987 and AAP82989.
XX
SQ Sequence 29 AA;
XX
Query Match 23.2%; Score 76; DB 9; Length 29;
Best Local Similarity 48.3%; Pred. No. 0.011;
Matches 14; Conservative 6; Mismatches 9; Indels 0; Gaps 0;
QY 10 FSQIYGVVYGVNTHFVPSNVPLKEVLWK 38
||| ||| : ||||| : | ||| :
Db 1 FSQDIYGAMNGSVTFYVSESPFTEIMKG 29
||| ||| : ||||| : | ||| :

RESULT 10
AAP82986
ID AAP82986 standard; protein; 29 AA.
XX
AC AAP82986;
XX
DT 06-DEC-1990 (first entry)
XX
DE Sheep erythrocyte target structure for T lymphocytes #2.
XX erythrocyte target structure for T lymphocytes; (TILTS);
KW immune disease; immunomodulator.
XX
OS Ovis aries.
XX
PN EP268995-A.
XX
PD 01-JUN-1988.
XX
PF 19-NOV-1987; 87EP-0008695.
XX
PR 22-NOV-1986; 86DE-3639920.
XX
PA (FARH ) HOECHST AG.
XX
PI Hunig T, Tiefenthaler G, Mitnacht R, Meuer S;
XX
DR WPI; 1988-148827/22.
XX
PT Homogeneous sheep TILTS and human TILTS glycoprotein - useful as
XX immunomodulators and for monoclonal antibody prodn.
XX
PS Claim 1; Page 9; 9pp; German.
XX
CC Specific example of N-terminal sequence of sheep TILTS. The entire
XX TILTS protein is used for immunising an animal and antibodies are
CC isolated from the serum. These antibodies are in turn used to react
CC with human TILTS glycoprotein in an immunoassay. Both the sheep and
CC human derived proteins are useful as immunomodulators.
CC See also AAP80286 and AAP82987-P82989.
XX

```

```

XX
SQ Sequence 29 AA;
XX
Query Match 22.3%; Score 73; DB 9; Length 29;
Best Local Similarity 46.4%; Pred. No. 0.027;
Matches 13; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
QY 11 SQIYGVVYGVNTHFVPSNVPLKEVLWK 38
||| ||| : ||||| : | ||| :
Db 2 SQDIYGAMNGSVTFYVSESPFTEIMFK 29
||| ||| : ||||| : | ||| :

RESULT 11
AAP82989
ID AAP82989 standard; protein; 29 AA.
XX
AC AAP82989;
XX
DT 06-DEC-1990 (first entry)
XX
DE Sheep erythrocyte target structure for T lymphocytes #5.
XX erythrocyte target structure for T lymphocytes; (TILTS);
KW immune disease; immunomodulator.
XX
OS Ovis aries.
XX
PN EP268995-A.
XX
PD 01-JUN-1988.
XX
PF 19-NOV-1987; 87EP-0008695.
XX
PR 22-NOV-1986; 86DE-3639920.
XX
PA (FARH ) HOECHST AG.
XX
PI Hunig T, Tiefenthaler G, Mitnacht R, Meuer S;
XX
DR WPI; 1988-148827/22.
XX
PT Homogeneous sheep TILTS and human TILTS glycoprotein - useful as
XX immunomodulators and for monoclonal antibody prodn.
XX
PS Claim 1; Page 9; 9pp; German.
XX
CC Specific example of N-terminal sequence of sheep TILTS. The entire
XX TILTS protein is used for immunising an animal and antibodies are
CC isolated from the serum. These antibodies are in turn used to react
CC with human TILTS glycoprotein in an immunoassay. Both the sheep and
CC human derived proteins are useful as immunomodulators.
CC See also AAP80286, AAP82986-P82988.
XX
SQ Sequence 29 AA;
XX
Query Match 21.7%; Score 71; DB 9; Length 29;
Best Local Similarity 44.8%; Pred. No. 0.049;
Matches 13; Conservative 6; Mismatches 10; Indels 0; Gaps 0;
QY 10 FSQIYGVVYGVNTHFVPSNVPLKEVLWK 38
||| ||| : ||||| : | ||| :
Db 1 FSSDIYGAMNGSVTFYVSESPFTEIMLK 29
||| ||| : ||||| : | ||| :

RESULT 12
AAP82987
ID AAP82987 standard; protein; 29 AA.
XX
AC AAP82987;
XX
DT 06-DEC-1990 (first entry)
XX
DE Sheep erythrocyte target structure for T lymphocytes #3.
XX

```

XX erythrocyte target structure for T lymphocytes; (Tilts);
 KW immune disease; immunomodulator.

XX Ovis aries.

XX EP268995-A.

XX 01-JUN-1988.

XX 19-NOV-1987; 87EP-0008695.

XX 22-NOV-1986; 86DE-3639920.

XX (FARH) HOECHST AG.

XX Hunig T, Tiefenthaler G, Mitnacht R, Meuer S;

XX WPI; 1988-148827/22.

XX Homogeneous sheep Tilts and human Tilts glycoprotein - useful as
 PT immunomodulators and for monoclonal antibody prodn.

XX Claim 1; Page 9; 9pp; German.

XX Specific example of N-terminal sequence of sheep Tilts. The entire
 CC Tilts protein is used for immunising an animal and antibodies are
 CC isolated from the serum. These antibodies are in turn used to react
 CC with human Tilts glycoprotein in an immunoassay. Both the sheep and
 CC human derived proteins are useful as immunomodulators.
 CC See also AAP80286, AAP82986 and AAP82988-P82989.

XX Sequence 29 AA;

Query Match 19.6%; Score 64; DB 9; Length 29;
 Best Local Similarity 42.9%; Pred. No. 0.39; Mismatches 10; Indels 0; Gaps 0;
 Matches 12; Conservative 6;

QY 11 SQQIYGVVYGNVTFHVPSPNPKLEVLWK 38

DB 2 SSDIYGAMNGSVTFYVSESQPFTEIMAK 29

RESULT 13

AAP80286
 ID AAP80286 standard; protein; 29 AA.

XX AAP80286;

XX 06-DEC-1990 (first entry)

XX Sheep erythrocyte target structure for T lymphocytes #1.

XX erythrocyte target structure for T lymphocytes; (Tilts);
 KW immune disease.

XX Ovis aries.

XX Key Location/Qualifiers

FT Misc-difference 1 /label= V, F

FT Misc-difference 3 /label= Q, S

FT EP268995-A.

XX 01-JUN-1988.

XX 19-NOV-1987; 87EP-0008695.

XX 22-NOV-1986; 86DE-3639920.

XX (FARH) HOECHST AG.

XX Hunig T, Tiefenthaler G, Mitnacht R, Meuer S;
 XX WPI; 1988-148827/22.
 XX Homogeneous sheep Tilts and human Tilts glycoprotein - useful as
 PT immunomodulators and for monoclonal antibody prodn.

XX Claim 1; Page 9; 9pp; German.

XX Generic N-terminal sequence of sheep Tilts. The entire Tilts
 CC protein is used for immunising an animal and antibodies are
 CC isolated from the serum. These antibodies are in turn used to react
 CC with human Tilts glycoprotein in an immunoassay. Both the sheep and
 CC human derived proteins are useful as immunomodulators.

XX The residue at posn. 12 can be any amino acid but is preferably
 CC Ser.

XX See also AAP82986-P82989.

XX Sequence 29 AA;

Query Match 19.0%; Score 62; DB 9; Length 29;
 Best Local Similarity 42.9%; Pred. No. 0.71;
 Matches 12; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 11 SQQIYGVVYGNVTFHVPSPNPKLEVLWK 38

DB 2 SXDIYGAMNGXVTFYVSESQPFTEIMXK 29

RESULT 14

AAP60315
 ID AAP60315 standard; Protein; 29 AA.

XX AAP60315;

XX 27-FEB-1995 (first entry)

XX Sheep LFA-3 N-terminal peptide.

XX Mature form; human; ovine; LFA-3; antiviral agent; virus; infection;
 KW propagation cycle; HIV; PCR; primer; sheep; homologue; D1D2 protein.

XX Ovis sp.

XX Key Location/Qualifiers

FT Misc-difference 1 /label= Val, Phe

FT Misc-difference 3 /label= Gln, Ser

FT Misc-difference 12 /note= "Unidentified amino acid"

FT Misc-difference 28 /note= "Unidentified amino acid"

XX JP06157334-A.

XX 03-JUN-1994.

XX 27-NOV-1992; 92JP-0318934.

XX 27-NOV-1992; 92JP-0318934.

XX (KANF) KANEBUCHI KAGAKU KOGYO KK.

XX WPI; 1994-221795/27.

XX An antiviral agent containing LFA-3 - useful on a virus with the
 PT same infection propagation cycle as HIV.

XX Disclosure; Page 14; 18pp; Japanese.

XX This sequence represents the N-terminal fragment of ovine LFA-3. This
 CC

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 09:00:21 ; Search time 11 Seconds
(without alignments)
111.899 Million cell updates/sec

Title: US-09-730-465-2_COPY_20_80

perfect score: 327
Sequence: 1 LHCFGFISCFSSQIIYGVVYG.....DKVAELENSEFRAFSSFNRR 61

Scoring table: BLOSUM62

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 61208

Minimum DB seq length: 0

Maximum DB seq length: 61

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

```
Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/us08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	50.5	15.4	56	10	US-09-864-761-33712		Sequence 33712, A
2	50.5	15.4	56	10	US-09-864-761-34125		Sequence 34125, A
3	48	14.7	57	10	US-09-764-869-1235		Sequence 1235, Ap
4	47	14.4	51	10	US-09-799-514-19		Sequence 19, Appl
5	46	14.1	59	10	US-09-864-761-39695		Sequence 39695, A
6	43.5	13.3	33	10	US-09-864-761-34725		Sequence 34725, A
7	43.5	13.3	46	10	US-09-764-869-641		Sequence 641, App
8	43.5	13.3	57	10	US-09-864-761-33732		Sequence 33732, A
9	43	13.1	34	10	US-09-815-242-10873		Sequence 10873, A
10	42.5	13.0	49	9	US-10-001-835-147		Sequence 147, App
11	42.5	13.0	59	10	US-09-864-761-42851		Sequence 42851, A
12	42	12.8	34	10	US-09-864-761-33625		Sequence 33625, A
13	42	12.8	34	10	US-09-864-761-34165		Sequence 34165, A
14	42	12.8	35	10	US-09-864-761-33626		Sequence 33626, A
15	42	12.8	57	9	US-09-796-692-2396		Sequence 2396, Ap
16	41.5	12.7	58	10	US-09-864-761-49002		Sequence 49002, A
17	41	12.5	28	10	US-09-864-761-46128		Sequence 46128, A
18	41	12.5	51	10	US-09-864-761-34679		Sequence 34679, A
19	40.5	12.4	41	10	US-09-997-498-2		Sequence 2, Appli

Sequence 34, Appl
Sequence 36, Appl
Sequence 37, Appl
Sequence 38, Appl
Sequence 38, Appl
Sequence 37, Appl
Sequence 38, Appl
Sequence 36, Appl
Sequence 37, Appl
Sequence 38, Appl
Sequence 45809, A
Sequence 42859, A
Sequence 34013, A
Sequence 4, Appl
Sequence 34123, A
Sequence 1220, Ap
Sequence 33, Appl
Sequence 36347, A
Sequence 38284, A
Sequence 182, App
Sequence 5, Appl
Sequence 3, Appl
Sequence 8, Appl
Sequence 37982, A
Sequence 1417, Ap
Sequence 35796, A

ALIGNMENTS

```

RESULT 1
US-09-864-761-33712
: Sequence 33712, Application US/09864761
: Patent NO. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
: TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
: FILE REFERENCE: Aemica-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: PRIOR FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30

```

```
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33712
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000087.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 6.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.9
; OTHER INFORMATION: SWISSPROT HIT: P18859, EVALUE 5.00e-24
; OTHER INFORMATION: EST_HUMAN HIT: AA128761.1, EVALUE 5.00e-24
US-09-864-761-33712
```

```
Query Match 15.4%; Score 50.5; DB 10; Length 56;
Best Local Similarity 24.1%; Pred. No. 7.2;
Matches 13; Conservative 14; Mismatches 24; Indels 3; Gaps 1;
```

```
QY 7 ISCFSSQIY---GVVYGNVTFHVPNSVPLKVKKKQKDKVAELENSEFRAPSS 57
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1 ISMILQRLFRFSSVIRSAVSVHLRRNIGVTAVAFNKELDPQKLFVDKIREYKS 54
```

RESULT 2

```
US-09-864-761-34125
; Sequence 34125, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34125
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000139.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 5.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
; OTHER INFORMATION: EST_HUMAN HIT: AA128761.1, EVALUE 5.00e-24
; OTHER INFORMATION: SWISSPROT HIT: P18859, EVALUE 5.00e-24
US-09-864-761-34125
```

```
Query Match 15.4%; Score 50.5; DB 10; Length 56;
Best Local Similarity 24.1%; Pred. No. 7.2;
Matches 13; Conservative 14; Mismatches 24; Indels 3; Gaps 1;
```

```
QY 7 ISCFSSQIY---GVVYGNVTFHVPNSVPLKVKKKQKDKVAELENSEFRAPSS 57
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1 ISMILQRLFRFSSVIRSAVSVHLRRNIGVTAVAFNKELDPQKLFVDKIREYKS 54
```

RESULT 3

```
US-09-764-869-1235
; Sequence 1235, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1235
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (26)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (28)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (49)
```

;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 34165
;; LENGTH: 34
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AP000085.1
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.7
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.8
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.9
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2
US-09-864-761-34165

Query Match 12.8%; Score 42; DB 10; Length 34;
Best Local Similarity 46.7%; Pred. No. 47;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 35 VLWKKQKDKVAELEN 49
:|||: :|||:
Db 2 ILWNKQKQDSQLES 16

RESULT 14

US-09-864-761-33626
;; Sequence 33626, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
;; FILE REFERENCE: Aecomica-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 33626
;; LENGTH: 35
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AP000223.1
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.9
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.2
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.1
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.7
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.3
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.7
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.5
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.3
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3
US-09-864-761-33626

Query Match 12.8%; Score 42; DB 10; Length 35;
Best Local Similarity 46.7%; Pred. No. 48;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 35 VLWKKQKDKVAELEN 49
:|||: :|||:
Db 2 ILWNKQKQDSQLES 16

RESULT 15

US-09-796-692-2396
;; Sequence 2396, Application US/09796692
;; Publication No. US20020198362A1
;; GENERAL INFORMATION:
;; APPLICANT: Gaiger, Alexander
;; APPLICANT: Algate, Paul A.
;; APPLICANT: Mannion, Jane
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
;; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
;; FILE REFERENCE: 2077.001200
;; CURRENT APPLICATION NUMBER: US/09/796,692
;; CURRENT FILING DATE: 2001-03-01
;; PRIOR APPLICATION NUMBER: 60/186,126
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: 60/190,479
;; PRIOR FILING DATE: 2000-03-17
;; PRIOR APPLICATION NUMBER: 60/200,545
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: 60/200,303
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,779

```
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2396
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-796-692-2396

Query Match      12.8%; Score 42; DB 9; Length 57;
Best Local Similarity 45.0%; Pred. No. 88;
Matches 9; Conservative 6; Mismatches 3; Indels 2; Gaps 2;

Qy  15 YGVVYGNVTFTVPS-NVPLK 33
    |||||:|:|:|:|
Db   37 YG-IYGNSTYHLPLNFSLNK 55
```

Search completed: January 28, 2003, 09:04:50
Job time : 12 secs


```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34725
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007115.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 18
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 11
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.95
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 34
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.3
; OTHER INFORMATION: SWISSPROT HIT: Q9XSD9, EVALUE 7.00e-06
; OTHER INFORMATION: EST_HUMAN HIT: AV753066.1, EVALUE 5.00e-05
US-09-864-761-34725
```

```
Query Match          13.3%; Score 43.5; DB 10; Length 33;
Best Local Similarity 26.3%; Pred. No. 29;
Matches 9; Conservative 10; Mismatches 12; Indels 3; Gaps 1;
```

```
QY 27 PSNVPLKEVLWKQKDKVAELENSEFRASFKN 60
      |::| | | | | | | | | | | | | | | |
Db 1 PKDLPDPTDLLDQNNKITIKGDGFK---NLKN 31
```

```
RESULT 7
US-09-764-869-641
; Sequence 641, Application US/09764869
; Patent No. US20020061521a1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 641
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: SITE
; LOCATION: (37)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-869-641

Query Match          13.3%; Score 43.5; DB 10; Length 46;
Best Local Similarity 38.5%; Pred. No. 44;
Matches 10; Conservative 3; Mismatches 10; Indels 3; Gaps 1;

QY 15 YGVVYGNVTE---HVPSPNVLKVELW 37
      ||:| | | | | | | | | | | |
Db 4 YGIWNNILLPVGHSPNNVENTYKLM 29
```

```
RESULT 8
US-09-864-761-33732
; Sequence 33732, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeo mica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33732
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL050337.6
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.3
```



```
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 42851
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006452.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EST_HUMAN HIT: BE738056.1, EVALUAE 3.60e+00
US-09-864-761-42851
```

```
Query Match 13.0%; Score 42.5; DB 10; Length 59;
Best Local Similarity 40.7%; Pred. No. 79;
Matches 11; Conservative 0; Mismatches 7; Indels 9; Gaps 1;
```

```
Qy 9 CFSQIYGVVYGVNTHFVPSNVPKEY 35
||| ||| ||| ||| |||
Db 11 CFPQQ-----PFHAPSGSSKEY 28
```

RESULT 12

```
US-09-864-761-33625
; Sequence 33625, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33625
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000085.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.3
US-09-864-761-33625
```

```
Query Match 12.8%; Score 42; DB 10; Length 34;
Best Local Similarity 46.7%; Pred. No. 47;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
```

```
Qy 35 VLWKKOKDKVAELEN 49
||| ||| ||| |||
Db 2 ILWKKQKQDDSQLES 16
```

RESULT 13

```
US-09-864-761-34165
; Sequence 34165, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
```



APPLICANT: Murphy, Randall B.
APPLICANT: Schuster, David I.
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/118,270
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: MURPHY-2A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 307:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-118-270-307

Query Match 14.1%; Score 46; DB 1; Length 45;
Best Local Similarity 57.1%; Pred. No. 34;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 14 IYGVVYGNVTFHVP 27
:||:| |||:
Db 5 VYGLVDGLVTFYLP 18

Search completed: January 28, 2003, 09:01:40
Job time : 15 secs

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/118,270
; FILING DATE: 09-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: MURPHY=2A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 306:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 44 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-118-270-306

Query Match 14.1%; Score 46; DB 1; Length 44;
Best Local Similarity 57.1%; Pred. No. 33;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 14 IYGVVYGVNVTFRVP 27
Db 4 VYGLVDGLVTFYLP 17
:|:|:|:|:|:|:|

RESULT 13
PCT-US93-08528-305
; SEQUENCE 305, Application PC/TUS9308528
; GENERAL INFORMATION:
; APPLICANT: New York University
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08528
; FILING DATE: 09-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: MURPHY=2 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 306:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 44 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US93-08528-306

Query Match 14.1%; Score 46; DB 5; Length 44;
Best Local Similarity 57.1%; Pred. No. 33;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 14 IYGVVYGVNVTFRVP 27
Db 4 VYGLVDGLVTFYLP 17
:|:|:|:|:|:|:|

RESULT 15
US-08-118-270-307
; SEQUENCE 307, Application US/08118270
; Patent No. 5508384
; GENERAL INFORMATION:
; SEQUENCE CHARACTERISTICS:
```

```
;
; LENGTH: 44 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US93-08528-305

Query Match 14.1%; Score 46; DB 5; Length 44;
Best Local Similarity 57.1%; Pred. No. 33;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 14 IYGVVYGVNVTFRVP 27
Db 4 VYGLVDGLVTFYLP 17
:|:|:|:|:|:|:|

RESULT 14
PCT-US93-08528-306
; SEQUENCE 306, Application PC/TUS9308528
; GENERAL INFORMATION:
; APPLICANT: New York University
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08528
; FILING DATE: 09-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: MURPHY=2 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 306:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 44 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US93-08528-306

Query Match 14.1%; Score 46; DB 5; Length 44;
Best Local Similarity 57.1%; Pred. No. 33;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 14 IYGVVYGVNVTFRVP 27
Db 4 VYGLVDGLVTFYLP 17
:|:|:|:|:|:|:|

RESULT 15
US-08-118-270-307
; SEQUENCE 307, Application US/08118270
; Patent No. 5508384
; GENERAL INFORMATION:
```

;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; FRAGMENT TYPE: internal
;; ORIGINAL SOURCE:
US-08-103-170-26

Query Match 14.1%; Score 48; DB 2; Length 24;
Best Local Similarity 53.3%; Pred. No. 9;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 13 QYGVVYGVNVTFFHP 27
:|:|:| | | | | | |
Db 2 EVGLVDGLVTFYLP 16

RESULT 10
US-08-103-170-27
; Sequence 27, Application US/08103170
; Patent No. 5885824
; GENERAL INFORMATION:
; APPLICANT: Yamada, Tadataka
; TITLE OF INVENTION: Recombinant Genomic Clones Encoding
; TITLE OF INVENTION: Histamine H1, H2, and H3 Receptors, Methods For Production
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202

;; COMPUTER READABLE FORM: disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/103,170
;; FILING DATE:
;; CLASSIFICATION: 435

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/633,060
;; FILING DATE: 24-DEC-1990

;; ATTORNEY/AGENT INFORMATION:
;; NAME: Lavalleye, Jean-Paul
;; REGISTRATION NUMBER: 31,451
;; REFERENCE/DOCKET NUMBER: 2363-017-55

;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703)521-4500
;; TELEFAX: (703)486-2347

;; INFORMATION FOR SEQ ID NO: 27:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 24 amino acids
;; TYPE: amino acid

;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; FRAGMENT TYPE: internal
;; ORIGINAL SOURCE:
;; ORGANISM: Dog

US-08-103-170-27
Query Match 14.1%; Score 46; DB 2; Length 24;
Best Local Similarity 57.1%; Pred. No. 16;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 14 IYGVVYGVNVTFFHP 27
:|:|:| | | | | | |
Db 3 YVGLVDGLVTFYLP 16

RESULT 11
US-08-118-270-305
; Sequence 305, Application US/08118270
; Patent No. 5508384
; GENERAL INFORMATION:
; APPLICANT: Murphy, Randall B.
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/118,270
;; FILING DATE: 09-SEP-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/943,236
;; FILING DATE: 10-SEP-1992

;; ATTORNEY/AGENT INFORMATION:
;; NAME: Townsend, Kevin G.
;; REGISTRATION NUMBER: 34,033

;; REFERENCE/DOCKET NUMBER: MURPHY-2A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-628-5197
;; TELEFAX: 202-737-3528

;; TELEX: 248633
;; INFORMATION FOR SEQ ID NO: 305:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 44 amino acids
;; TYPE: amino acid

;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-118-270-305

Query Match 14.1%; Score 46; DB 1; Length 44;
Best Local Similarity 57.1%; Pred. No. 33;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 14 IYGVVYGVNVTFFHP 27
:|:|:| | | | | | |
Db 4 YVGLVDGLVTFYLP 17

RESULT 12
US-08-118-270-306
; Sequence 306, Application US/08118270
; Patent No. 5508384
; GENERAL INFORMATION:
; APPLICANT: Murphy, Randall B.
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004


```

: STREET: 875 Third Avenue
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10022-6250
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US92/02050
: FILING DATE: 19920312
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/667,971
: FILING DATE: 12-MAR-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/770,967
: FILING DATE: 07-OCT-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: HALEY, James F., Jr.
: REGISTRATION NUMBER: 27,794
: REFERENCE/DOCKET NUMBER: B15ICIP2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212)715-0600
: TELEFAX: (212)715-0673
: TELEX: 14-8367
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 50 amino acids
: TYPE: AMINO ACID
: STRANDEDNESS: single
: TOPOLOGY: linear
: HYPOTHEtical: NO
: ANTI-SENSE: NO
: PCT-US92-02050-2

Query Match 79.5%; Score 260; DB 5; Length 50;
Best Local Similarity 100.0%; Pred. No. 6.5e-27; Indels
Matches 50; Conservative 0; Mismatches 0;

QY 10 FSQQIYGVVYGVNTHFVPSNVPLKEVLKKQDKVAELENSEFRAFSSFK 59
|||||
Db 1 FSQQIYGVVYGVNTHFVPSNVPLKEVLKKQDKVAELENSEFRAFSSFK 38
|||||

RESULT 6
5185441-1
: Patent No. 5185441
: APPLICANT: WALLNER, BARBARA P.; HESSTONS, CATHERINE
: TITLE OF INVENTION: DNA SEQUENCES, RECOMBINANT DNA
: MOLECULES AND PROCESSES FOR PRODUCING PI-LINKED LYPHCYTE
: FUNCTION ASSOCIATED ANTIGEN-3
: NUMBER OF SEQUENCES: 41
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/237,309
: FILING DATE: 26-AUG-1988
: SEQ ID NO: 1:
: LENGTH: 38
: 5185441-1

Query Match 61.2%; Score 200; DB 6; Length 38;
Best Local Similarity 100.0%; Pred. No. 2.9e-19; Indels
Matches 38; Conservative 0; Mismatches 0;

QY 10 FSQQIYGVVYGVNTHFVPSNVPLKEVLKKQDKVAEL 47
|||||
Db 1 FSQQIYGVVYGVNTHFVPSNVPLKEVLKKQDKVAEL 38
|||||

RESULT 7
SS-08-328-152A-3

```

;
; TOPOLOGY: linear
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-07-940-861-2

Query Match 79.5%; Score 260; DB 1; Length 50;
Best Local Similarity 100.0%; Pred. No. 6.5e-27;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 FSQIYGVVYGVNVTFHVPSNVPLKEVLWKQKDKVAELENSEFRAFSSFK 59
|||||
Db 1 FSQIYGVVYGVNVTFHVPSNVPLKEVLWKQKDKVAELENSEFRAFSSFK 50

RESULT 2
US-08-459-512-2
; Sequence 2, Application US/08459512
; Patent No. 5728677
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: WALLNER, Barbara P.
; APPLICANT: MILLER, Glenn T.
; APPLICANT: ROSA, Margaret D.
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
; TITLE OF INVENTION: FUNCTION ASSOCIATED ANTIGEN 3
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 875 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022-6250
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,512
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02050
; FILING DATE: 12-MAR-1992
; APPLICATION NUMBER: US 07/667,971
; FILING DATE: 12-MAR-1991

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-459-512-2

Query Match 79.5%; Score 260; DB 1; Length 50;
Best Local Similarity 100.0%; Pred. No. 6.5e-27;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10 FSQIYGVVYGVNVTFHVPSNVPLKEVLWKQKDKVAELENSEFRAFSSFK 59
|||||
Db 1 FSQIYGVVYGVNVTFHVPSNVPLKEVLWKQKDKVAELENSEFRAFSSFK 50

RESULT 4
US-08-460-132-2
; Sequence 2, Application US/08460132
; Patent No. 5928643
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.

Db 1 FSQIYGVVYGVNVTFHVPSNVPLKEVLWKQKDKVAELENSEFRAFSSFK 50
|||||

RESULT 3
US-08-459-657-2
; Sequence 2, Application US/08459657
; Patent No. 5914111
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: WALLNER, Barbara P.
; APPLICANT: MILLER, Glenn T.
; APPLICANT: ROSA, Margaret D.
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
; TITLE OF INVENTION: FUNCTION ASSOCIATED ANTIGEN 3
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 875 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022-6250
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,657
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02050
; FILING DATE: 12-MAR-1992
; APPLICATION NUMBER: US 07/667,971
; FILING DATE: 12-MAR-1991

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-459-657-2

Query Match 79.5%; Score 260; DB 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 6.5e-27;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10 FSQIYGVVYGVNVTFHVPSNVPLKEVLWKQKDKVAELENSEFRAFSSFK 59
|||||

Db 1 FSQIYGVVYGVNVTFHVPSNVPLKEVLWKQKDKVAELENSEFRAFSSFK 50
|||||

RESULT 4
US-08-460-132-2
; Sequence 2, Application US/08460132
; Patent No. 5928643
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 08:58:45 ; Search time 14 Seconds
(without alignments)
128.200 Million cell updates/sec

Title: US-09-730-465-2_COPY_20_80

Perfect score: 327

Sequence: 1 LHCFGFCISFQQIYGVVYG.....DKVAELENSEFFRAFSFKNR 61

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 186412

Minimum DB seq length: 0

Maximum DB seq length: 61

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	260	79.5	50	1	US-07-940-861-2
2	260	79.5	50	1	US-08-459-512-2
3	260	79.5	50	2	US-08-459-657-2
4	260	79.5	50	2	US-08-460-132-2
5	260	79.5	50	5	PCT-US92-02050-2
6	200	61.2	38	6	5185441-1
7	62	19.0	29	1	US-08-328-152A-3
8	54	16.5	11	6	5185441-2
9	48	14.7	24	2	US-08-103-170-26
10	46	14.1	24	2	US-08-103-170-27
11	46	14.1	44	1	US-08-118-270-305
12	46	14.1	44	1	US-08-118-270-306
13	46	14.1	44	5	PCT-US93-08528-305
14	46	14.1	44	5	PCT-US93-08528-306
15	46	14.1	45	1	US-08-118-270-307
16	46	14.1	45	5	PCT-US93-08528-307
17	46	14.1	47	1	US-08-118-270-308
18	46	14.1	47	5	PCT-US93-08528-308
19	42	12.8	7	6	5185441-12
20	40.5	12.4	41	1	US-08-865-773-5
21	40.5	12.4	41	4	US-09-400-716-2
22	40.5	12.4	41	4	US-09-424-127-5
23	40.5	12.4	41	6	5177060-2
24	40	12.2	43	4	US-09-015-030-6
25	39.5	12.1	41	1	US-07-709-091-1
26	39.5	12.1	41	1	US-07-709-091-4
27	39.5	12.1	41	1	US-07-709-091-5

28	39.5	12.1	41	1	US-07-709-091-6
29	39.5	12.1	41	1	US-07-715-752A-1
30	39.5	12.1	41	1	US-07-766-682A-4
31	39.5	12.1	41	1	US-08-104-862-1
32	39.5	12.1	41	1	US-08-104-862-4
33	39.5	12.1	41	1	US-08-104-862-5
34	39.5	12.1	41	1	US-08-104-862-6
35	39.5	12.1	41	1	US-08-162-178-1
36	39.5	12.1	41	1	US-08-865-773-1
37	39.5	12.1	41	4	US-08-981-189B-1
38	39.5	12.1	41	4	US-09-260-846-22
39	39.5	12.1	41	4	US-09-400-716-4
40	39.5	12.1	41	4	US-09-424-127-1
41	39.5	12.1	41	5	PCT-US92-05101-1
42	39.5	12.1	41	6	5177060-4
43	39.5	12.1	47	4	US-09-149-476-472
44	39.5	12.1	60	2	US-08-117-952-787
45	39	11.9	58	3	US-08-676-124-48

ALIGNMENTS

RESULT 1
US-07-940-861-2
; Sequence 2, Application US/07940861
; Patent No. 5547853
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: WALLNER, Barbara P.
; APPLICANT: MILLER, Glenn T.
; APPLICANT: ROSA, Margaret D.
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 875 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022-6250
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/940.861
; FILING DATE: 21-OCT-1992
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02050
; FILING DATE: 12-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/667,971
; FILING DATE: 12-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,967
; FILING DATE: 07-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY, James F., Jr.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B151CIP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)715-0600
; TELEFAX: (212)715-0673
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

Sequence 6, Appli
Sequence 1, Appli
Sequence 4, Appli
Sequence 1, Appli
Sequence 4, Appli
Sequence 5, Appli
Sequence 6, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 22, Appli
Sequence 4, Appli
Sequence 1, Appli
Patent No. 5177060
Sequence 472, App
Sequence 787, App
Sequence 48, Appli

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: January 28, 2003, 08:38:45 ; Search time 27.2194 Seconds
(without alignments)
1223.856 Million cell updates/sec

Title: US-09-730-465-2

Perfect score: 1326

Sequence: 1 MWAGSDAGALGVLSVVCLL.....VLYMNGILKCDKRPDRTNSN 250

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	1326	100.0	250	9 AAP81507	Human LFA-3 expres
2	1326	100.0	250	13 AAR27161	Transmembrane LFA-
3	1326	100.0	250	15 AAR64271	Human LFA-3. Homo
4	1326	100.0	250	17 AAW04370	Human cell adhesio
5	1326	100.0	250	21 AAY83133	Human transmembran
6	1326	100.0	250	22 AAB61157	Human transmembran
7	1326	100.0	250	23 AAU76225	Lymphocyte functio
8	1313	99.0	250	14 AAR34371	Sequence encoded b
9	1307	98.6	250	14 AAR34221	Amino acid sequenc
10	1289	97.2	250	13 AAR28365	Human LFA-3 protei

11	1250	94.3	240	11 AAR07604	Lymphocyte functio
12	1250	94.3	240	13 AAR27162	PI-linked LFA-3 us
13	1250	94.3	240	21 AAY83134	PI-linked human tr
14	1250	94.3	240	22 AAB61158	Human PI-linked tr
15	1250	94.3	240	23 AAU76226	PI-linked lymphocy
16	1247	94.0	240	14 AAR34222	Amino acid sequenc
17	1247	94.0	240	14 AAR34372	Sequence encoded b
18	1244	93.8	240	11 AAR05572	PI-linked LFA-3 en
19	1243	93.7	237	13 AAR20804	Human LFA-3 antige
20	1243	93.7	237	17 AAR91432	Human LFA-3 antige
21	1243	93.7	237	21 AAY96127	Human cell surface
22	1243	93.7	237	22 AAU02436	Human lymphocyte f
23	1237	93.3	237	18 AAW16687	Human CD58 GPI. H
24	1232	92.9	237	12 AAR14182	Human LFA-3(CD58)
25	1231	92.8	237	19 AAW80441	Human LFA-3 antige
26	1231	92.8	237	20 AAW86189	Human LFA-3 antige
27	1179	88.9	222	15 AAR60313	Human mature LFA-3
28	1094.5	82.5	280	22 ABG18595	Novel human diagno
29	1026.5	77.4	197	15 AAR60310	Human LFA-3 D1D2HC
30	647	48.8	134	13 AAR28364	Human LFA-3 D2 reg
31	639	48.2	135	13 AAR28369	Human LFA-3-delta-
32	630	47.5	347	13 AAR27163	CD2 binding LFA-3-
33	630	47.5	347	21 AAY83136	Human transmembran
34	630	47.5	347	22 AAB61160	Human LFA3TIP fusi
35	630	47.5	347	23 AAU76228	LFA3/IgG fusion pr
36	627	47.3	128	17 AAW04361	Human cell adhesio
37	627	47.3	133	17 AAW04362	Human cell adhesio
38	627	47.3	133	17 AAW04363	Human cell adhesio
39	623	47.0	120	13 AAR27157	LFA-3 CD2 binding
40	615	46.4	347	14 AAR34224	Sheep LFA-3. Ovis
41	471.5	35.6	225	15 AAR60314	Sheep LFA-3. Ovis
42	458.5	34.6	225	13 AAR28366	Sheep LFA-3 protei
43	447.5	33.7	199	13 AAR28367	LFA-3 TM region de
44	447.5	33.7	199	15 AAR60311	Sheep LFA-3 delta
45	440.5	33.2	199	15 AAR60316	Sheep LFA-3 delta

ALIGNMENTS

```

RESULT 1
AAP81507
ID AAP81507 standard; protein; 250 AA.
AC AAP81507;
XX
XX 22-OCT-1990 (first entry)
XX
XX Human LFA-3 expressed by cDNA in phage lambda HT16.
XX
XX Lymphocyte function associated antigen-3; adhesion inhibition;
XX T-lymphocytes; immune suppression.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Protein 29..222
XX FT /label=human LFA-3
XX FT Peptide 1..28
XX FT /label=signal peptide
XX
XX PN W080809820-A.
XX
XX PD 15-DEC-1988.
XX
XX PF 03-JUN-1988; 88WO-US01924.
XX
XX PR 03-JUN-1987; 87US-0057615.
XX
XX PA (BIOJ ) BIOGEN NV; (DANA- ) DANA-FARBER CANCER INST.
XX
XX PI Haley JF;
XX

```

DR WPI; 1988-368634/51.
 XX N-PSDB; AAN81956.
 XX DNA sequences encoding Lymphocyte Function Associated Antigen-3 -
 PT which inhibits adhesion between T-lymphocytes and target cells.
 XX
 PS Disclosure;); 46pp; English.
 XX The polypeptide binds to CD2, the surface receptor of T-lymphocytes.
 CC They also inhibit adhesion between T-lymphocytes and target cells
 CC and may be used to block or to augment the immune response. They
 CC interfere with the interaction of helper T-cells and antigen-presenting
 CC cells and may be used to target specific T-cells for lysis and immune
 CC suppression or to deliver drugs such as lymphokines to the specific
 CC target T-cells. Comparison of the sequence with the N-terminal from
 CC LFA-3 purified from human erythrocytes suggests that amino acid 1-28
 CC comprise a signal sequence and amino acid 29-222 is the mature LFA-3.
 CC See also AAN81956-N81958.
 XX
 XX Sequence 250 AA;
 SQ

Query Match 100.0%; Score 1326; DB 9; Length 250;
 Best Local Similarity 100.0%; Pred. No. 7.2e-118;
 Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWAGSDAGRALGVLVWCLLHCFGFISCFSQIYGVVGNVTFHVPNSVPLKEVLKKQK 60
 Db 1 MWAGSDAGRALGVLVWCLLHCFGFISCFSQIYGVVGNVTFHVPNSVPLKEVLKKQK 60
 QY 61 DKVAELENSEFRASFKNRVYLDTVSGSLTIYNTSSDEDEYEMESPNITDTMKFFLYV 120
 Db 61 DKVAELENSEFRASFKNRVYLDTVSGSLTIYNTSSDEDEYEMESPNITDTMKFFLYV 120
 QY 121 LESLPSPTLTALTNNGSIEVOCMIPEHYNHSHRGLIMYSWDCPMEOCKRNSTSIYFKMEND 180
 Db 121 LESLPSPTLTALTNNGSIEVOCMIPEHYNHSHRGLIMYSWDCPMEOCKRNSTSIYFKMEND 180
 QY 181 LPQKIOCTLSNPLFNTSSIIILTCIPSSGHSRHRALIPILAVITTCIVLYMNGILKC 240
 Db 181 LPQKIOCTLSNPLFNTSSIIILTCIPSSGHSRHRALIPILAVITTCIVLYMNGILKC 240
 QY 241 DRKPDRTNSN 250
 Db 241 DRKPDRTNSN 250

RESULT 2
 AAR27161
 ID AAR27161 standard; Protein; 250 AA.
 XX
 AC AAR27161;
 XX
 DT 20-MAY-1998 (first entry)
 XX
 XX Transmembrane LFA-3.
 XX
 KW lymphocyte associated antigen-3; T-lymphocyte accessory molecule;
 KW deletion mutant; CD2 binding site; immunomodulator.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT peptide 1..28
 FT /label= signal
 FT protein 29..250
 FT /label= transmembrane_LFA-3
 FT region 39..98
 FT /note= "region deleted in M100 mutant"
 FT region 99..158
 FT /note= "region deleted in M101 mutant"
 FT region 159..208
 FT /note= "region deleted in M102 mutant"
 XX
 XX ;

PN EP503648-A.
 XX
 PD 16-SEP-1992.
 XX
 PF 12-MAR-1992; 92EP-0104320.
 XX
 PR 12-MAR-1991; 91US-0667971.
 PR 07-OCT-1991; 91US-0770967.
 XX
 PA (BIOJ) BIOGEN INC.
 XX
 PI Miller GT, Rosa MD, Wallner BP;
 XX
 XX WPI; 1992-309760/38.
 DR N-PSDB; AAQ28677.
 XX
 XX CD2-binding domain of lymphocyte function associated antigen-3
 PT and DNA - for diagnosing and treating inflammation and
 PT auto-immune diseases, e.g. systemic lupus erythematosus and
 PT rheumatoid arthritis
 XX
 XX Example 7; Fig 6; 85pp; English.
 PS
 CC Transmembrane LFA-3 cDNA (previously sequenced) was used to
 CC generate a series of deletion mutants to map the CD2-binding
 CC domain. One of the initial series of 30 nucleotide deletions
 CC ("N57" in which nucleotides 244-273 are deleted) gave rise to an
 CC expression product which failed to bind to Jurkat cells expressing
 CC CD2. The mutant M57 protein has 10 amino acids adjacent to an
 CC N-glycosylation site removed. To investigate whether the region of
 CC the M57 mutation is directly involved in CD2-binding or indirectly
 CC disrupts binding by causing a conformational change, three larger
 CC deletions were made (M100-M102 - see Features Table). The M100
 CC deletion mutant did not bind to Jurkat cells; the other two mutants
 CC did. The entire CD2 binding domain is encompassed in the M100
 CC region.
 XX
 SQ Sequence 250 AA;
 Query Match 100.0%; Score 1326; DB 13; Length 250;
 Best Local Similarity 100.0%; Pred. No. 7.2e-118;
 Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWAGSDAGRALGVLVWCLLHCFGFISCFSQIYGVVGNVTFHVPNSVPLKEVLKKQK 60
 Db 1 MWAGSDAGRALGVLVWCLLHCFGFISCFSQIYGVVGNVTFHVPNSVPLKEVLKKQK 60
 QY 61 DKVAELENSEFRASFKNRVYLDTVSGSLTIYNTSSDEDEYEMESPNITDTMKFFLYV 120
 Db 61 DKVAELENSEFRASFKNRVYLDTVSGSLTIYNTSSDEDEYEMESPNITDTMKFFLYV 120
 QY 121 LESLPSPTLTALTNNGSIEVOCMIPEHYNHSHRGLIMYSWDCPMEOCKRNSTSIYFKMEND 180
 Db 121 LESLPSPTLTALTNNGSIEVOCMIPEHYNHSHRGLIMYSWDCPMEOCKRNSTSIYFKMEND 180
 QY 181 LPQKIOCTLSNPLFNTSSIIILTCIPSSGHSRHRALIPILAVITTCIVLYMNGILKC 240
 Db 181 LPQKIOCTLSNPLFNTSSIIILTCIPSSGHSRHRALIPILAVITTCIVLYMNGILKC 240
 QY 241 DRKPDRTNSN 250
 Db 241 DRKPDRTNSN 250

RESULT 3
 AAR64271
 ID AAR64271 standard; Protein; 250 AA.
 XX
 AC AAR64271;
 XX
 DT 27-FEB-1995 (first entry)
 XX
 XX Human LFA-3.

```
XX Mature form; human; ovine; LFA-3; antiviral agent; virus; infection;
KW propagation cycle; HIV; PCR; primer; sheep;homologue; DID2 protein.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..28
FT Protein /note= "Signal peptide"
FT /label= "Mature LFA-3"
FT /note= "Mature LFA-3"
XX
XX JF06157334-A.
XX
XX 03-JUN-1994.
XX
XX 27-NOV-1992; 92JP-0318934.
XX
XX 27-NOV-1992; 92JP-0318934.
XX
XX (KANF ) KANEBUCHI KAGAKU KOGYO KK.
XX
XX WPI; 1994-221795/27.
XX N-PSDB; AAQ70396.
XX
XX An antiviral agent containing LFA-3 - useful on a virus with the
XX same infection propagation cycle as HIV.
XX
XX Disclosure; Page 11-12; 18pp; Japanese.
XX
XX This sequence represents the full length form of human LFA-3. This
XX protein contains a signal peptide and the D1, D2, TM and HC domains.
XX LFA-3 is an antiviral agent which is effective on a virus which has the
XX same infection/propagation cycle as HIV. The cDNA encoding this
XX sequence was amplified by PCR. The antiviral activity of homologues of
XX this sequence was examined. The homologues were pref. deletion mutants
XX of this sequence, and could have one or more distinct domains removed.
XX (See also AAQ70396-413 and AAR60310-18).
XX
XX Sequence 250 AA;
XX
XX Query Match 100.0%; Score 1326; DB 15; Length 250;
XX Best Local Similarity 100.0%; Pred. No. 7.2e-118;
XX Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 MWAGSDAGRALGVLVWCLLHCFGFI SCFSQQIYGVVYGNVTFHVPVPLKEVLKKQK 60
XX Db 1 MWAGSDAGRALGVLVWCLLHCFGFI SCFSQQIYGVVYGNVTFHVPVPLKEVLKKQK 60
XX
XX 61 DKVAELENSEFRASFKNRVYLDVTSGSLTIYNTLTSSDEDEYEMESPNTDTMKFFLYV 120
XX Db 61 DKVAELENSEFRASFKNRVYLDVTSGSLTIYNTLTSSDEDEYEMESPNTDTMKFFLYV 120
XX
XX QY 121 LESLPSPTLTALTNGSTEVOCMIPSHYNSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
XX Db 121 LESLPSPTLTALTNGSTEVOCMIPSHYNSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
XX
XX QY 181 LPQKIQCTLSNPLFNTSSIIILTCIPSSGHSRHRALIPILAVITTCIVLYMNGILKC 240
XX Db 181 LPQKIQCTLSNPLFNTSSIIILTCIPSSGHSRHRALIPILAVITTCIVLYMNGILKC 240
XX
XX QY 241 DRKPDRTNSN 250
XX Db 241 DRKPDRTNSN 250
XX
XX RESULT 4
XX AAW04370
XX ID AAW04370 standard; Protein; 250 AA.
XX
XX AC AAW04370;
XX
XX DT 30-JUN-1997 (first entry)
```

```
XX Human cell adhesion protein LFA-3.
DE
XX
XX Human; T cell; MOLT-4; cell adhesion; LFA-3; immunosuppression;
KW cell line; immunosuppressant cell.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT Peptide 1..28
FT /label= sig_peptide
FT Peptide 29..250
FT /label= mat_peptide
XX
XX WO9633217-A1.
XX
XX 24-OCT-1996.
XX
XX 15-APR-1996; 96WO-JP01039.
XX
XX 27-DEC-1995; 95JP-0341959.
XX 19-APR-1995; 95JP-0094060.
XX 04-JUL-1995; 95JP-0169110.
XX
XX (KANF ) KANEGAFUCHI KAGAKU KOGYO KK.
XX
XX Fukuchi T, Kakutani T, Nishino T, Niwa H, Ohara T;
XX Okazaki T, Yamashita K;
XX
XX WPI; 1996-485732/48.
XX N-PSDB; AAT38555.
XX
XX Cell adhesion protein derived from LFA-3 - has immunosuppressant
XX effect and induces immunosuppressant cells when cultured in its
XX presence
XX
XX Example 2; Pages 63-64; 83pp; Japanese.
XX
XX The present sequence is the human T cell line MOLT-4 (ATCC
XX CRL-1582) derived cell adhesion protein LFA-3, from which a claimed
XX protein comprising human LFA-3 minus the first 28 amino-terminal
XX residues can be prepared. The protein has immunosuppressant
XX activity, and cells (e.g. T cells) cultured in its presence, and
XX preferably also in the presence of interleukin-2 and/or
XX granulocyte/macrophage colony stimulating factor, become
XX immunosuppressant cells with potent immunosuppressant, but low side
XX effects.
XX
XX Sequence 250 AA;
XX
XX Query Match 100.0%; Score 1326; DB 17; Length 250;
XX Best Local Similarity 100.0%; Pred. No. 7.2e-118;
XX Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 MWAGSDAGRALGVLVWCLLHCFGFI SCFSQQIYGVVYGNVTFHVPVPLKEVLKKQK 60
XX Db 1 MWAGSDAGRALGVLVWCLLHCFGFI SCFSQQIYGVVYGNVTFHVPVPLKEVLKKQK 60
XX
XX QY 61 DKVAELENSEFRASFKNRVYLDVTSGSLTIYNTLTSSDEDEYEMESPNTDTMKFFLYV 120
XX Db 61 DKVAELENSEFRASFKNRVYLDVTSGSLTIYNTLTSSDEDEYEMESPNTDTMKFFLYV 120
XX
XX QY 121 LESLPSPTLTALTNGSIEVQCMIPSHYNSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
XX Db 121 LESLPSPTLTALTNGSIEVQCMIPSHYNSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
XX
XX QY 181 LPQKIQCTLSNPLFNTSSIIILTCIPSSGHSRHRALIPILAVITTCIVLYMNGILKC 240
XX Db 181 LPQKIQCTLSNPLFNTSSIIILTCIPSSGHSRHRALIPILAVITTCIVLYMNGILKC 240
XX
XX QY 241 DRKPDRTNSN 250
XX Db 241 DRKPDRTNSN 250
```

```

RESULT 5
AAY83133
ID AAY83133 standard; Protein; 250 AA.
XX
AC AAY83133;
XX
DT 24-JUL-2000 (first entry)
XX
DE Human transmembrane LFA-3.
XX
KW LFA3; CD2; cell signalling; modulation; lymphocyte; T cell;
KW memory effector T lymphocyte; psoriatic arthritis;
KW rheumatoid arthritis; multiple sclerosis; atopic dermatitis;
KW uveitis; inflammatory bowel disease; Crohn's disease;
KW ulcerative colitis; cutaneous T cell lymphoma; inhibition; treatment;
KW therapy.
XX
OS Homo sapiens.
XX
PN W0200012113-A2.
XX
PD 09-MAR-2000.
XX
XX
XX 31-AUG-1999; 99WO-US20026.
XX
XX 31-AUG-1998; 98US-0098456.
XX
PA (BIOJ ) BIOGEN INC.
XX
PI Magilavy D;
XX
DR WPI; 2000-282928/24.
DR N-PSDB; AAZ93398.
XX
XX
XX Selective modulation of memory effector T lymphocytes by administration
XX of a CD2 binding agent which inhibits the CD2/LFA-3 interaction useful
XX for treating conditions such as inflammatory bowel diseases, psoriatic
XX arthritis
XX
PS Disclosure; Page 67-68; 76pp; English.
XX
XX Modulation of LFA3/CD2 interaction by administration of a CD2 binding
XX agent inhibits CD2 signalling and T cell proliferation and activation
XX and more particularly modulates the number and/or distribution of
XX memory effector T lymphocytes. The method can be used for treating a
XX condition in a subject where the condition is characterized by memory
XX effector T lymphocytes playing a role in the pathogenesis of the
XX condition such as psoriatic arthritis, rheumatoid arthritis,
XX multiple sclerosis, atopic dermatitis, uveitis, inflammatory bowel
XX disease, Crohn's disease, ulcerative colitis and cutaneous T cell
XX lymphoma and where the method comprises administering to the subject
XX an amount of CD2 binding agent sufficient to modulate the memory
XX effector T lymphocytes. The methods provide inhibition of antigen
XX specific interactions for all antigens present, inhibition of T cell
XX activation, no general immunosuppression, and possibly induction of
XX tolerance.
XX
SQ Sequence 250 AA;
Query Match 100.0%; Score 1326; DB 21; Length 250;
Best Local Similarity 100.0%; Pred. No. 7.2e-118;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MWAGSDAGRALGVLSVCLLHCFGFTSCFSQIQIGVYGVNVTFRVPSNVLKVKQK 60
DB 1 MWAGSDAGRALGVLSVCLLHCFGFTSCFSQIQIGVYGVNVTFRVPSNVLKVKQK 60
QY 61 DKVAELENSEPRAFSFKNRVYLDTVSGSLTIYNLTSSDEYEMESPNITDTMKFFLYV 120
DB 61 DKVAELENSEPRAFSFKNRVYLDTVSGSLTIYNLTSSDEYEMESPNITDTMKFFLYV 120

```

SQ Sequence 250 AA;

Query Match 100.0%; Score 1326; DB 22; Length 250;
 Best Local Similarity 100.0%; Pred. No. 7.2e-118; Mismatches 0; Indels 0; Gaps 0;
 Matches 250; Conservative 0;

Qy 1 MVAGSDAGRALGVLSVWCLLHCFGIFSCFSQQIYGVYGNVTFHVPNSVPLKEVLWKQK 60
 Db 1 MVAGSDAGRALGVLSVWCLLHCFGIFSCFSQQIYGVYGNVTFHVPNSVPLKEVLWKQK 60

Qy 61 DKVAELENSEFRAFSSFKNRVYLDTVSGSLTYNLTSSDEDEYEMESPNITDTMKFFLYV 120
 Db 61 DKVAELENSEFRAFSSFKNRVYLDTVSGSLTYNLTSSDEDEYEMESPNITDTMKFFLYV 120

Qy 121 LESLPSTLTCAITNGSIEVQCMPEHNSHRLIMYSWDCPMEQCKRNSTSIYFKMEND 180
 Db 121 LESLPSTLTCAITNGSIEVQCMPEHNSHRLIMYSWDCPMEQCKRNSTSIYFKMEND 180

Qy 181 LPQKIOCTLSNPLNTSSIIITTCIPSSGHSRHRYPALIPPLAVITTCIVLYMNGILKC 240
 Db 181 LPQKIOCTLSNPLNTSSIIITTCIPSSGHSRHRYPALIPPLAVITTCIVLYMNGILKC 240

Qy 241 DRKPDRTNSN 250
 Db 241 DRKPDRTNSN 250

RESULT 8
 AAR34371
 ID AAR34371 standard; Protein; 250 AA.
 AC AAR34371;
 XX
 XX
 DT 10-AUG-1993 (first entry)
 XX

XX PF 05-DEC-2000; 2000US-0730465.
 XX PR 07-OCT-1991; 91US-0770969.
 XX PR 02-APR-1992; 92US-0862022.
 XX PR 06-OCT-1992; 92WO-US08755.
 XX PR 06-JUN-1995; 95US-0466465.
 XX PA (BIOJ) BIOGEN INC.
 XX PI Wallner BP, Cooper KD;
 XX WPI: 2002-179118/23.
 XX DR N-PSDB; ABK15717.
 XX
 XX Treating conditions characterised by increased T cell activation and
 XX abnormal antigen presentation in the dermis and epidermis, e.g. atopic
 XX dermatitis, vitiligo and UV damage, by administering an inhibitor of
 XX the CD2/LFA-3 interaction -
 XX
 XX Claim 22; Page 14-15; 32pp; English.
 XX
 XX This invention relates to a novel method for preventing or treating skin
 XX conditions characterised by increased T cell activation and abnormal
 XX antigen presentation in the dermis and epidermis. The method comprises
 XX inhibition of the CD2/lymphocyte function associated antigen-3 (LFA-3)
 XX interaction by administering to a mammal, including a human, an
 XX inhibitor of the CD2/LFA-3 interaction. These inhibitors may be anti-
 XX LFA-3 antibody homologues, anti-CD2 antibody homologues, soluble LFA-3
 XX peptides, soluble CD-2 polypeptides, small molecules such as
 XX carbohydrates, LFA-3 and CD2 mimetic agents. This method is useful to
 XX treat skin disorders such as cell lymphoma, mycosis fungoides, allergic
 XX and irritant contact dermatitis, lichen planus, alopecia areata,
 XX pyoderma gangrenosum, vitiligo, ocular cicatricial pemphigoid,
 XX psoriasis, UV damage and urticaria. The present sequence represents the
 XX mammalian LFA-3 protein sequence used to create the inhibitors of
 XX CD-2/LFA interaction of the invention.
 XX
 XX Sequence 250 AA;
 XX
 XX Query Match 100.0%; Score 1326; DB 23; Length 250;
 XX Best Local Similarity 100.0%; Pred. No. 7.2e-118;
 XX Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVAGSDAGRALGVLSVWCLLHCFGIFSCFSQQIYGVYGNVTFHVPNSVPLKEVLWKQK 60
 Db 1 MVAGSDAGRALGVLSVWCLLHCFGIFSCFSQQIYGVYGNVTFHVPNSVPLKEVLWKQK 60

Qy 61 DKVAELENSEFRAFSSFKNRVYLDTVSGSLTYNLTSSDEDEYEMESPNITDTMKFFLYV 120
 Db 61 DKVAELENSEFRAFSSFKNRVYLDTVSGSLTYNLTSSDEDEYEMESPNITDTMKFFLYV 120

Qy 121 LESLPSTLTCAITNGSIEVQCMPEHNSHRLIMYSWDCPMEQCKRNSTSIYFKMEND 180
 Db 121 LESLPSTLTCAITNGSIEVQCMPEHNSHRLIMYSWDCPMEQCKRNSTSIYFKMEND 180

Qy 181 LPQKIOCTLSNPLNTSSIIITTCIPSSGHSRHRYPALIPPLAVITTCIVLYMNGILKC 240
 Db 181 LPQKIOCTLSNPLNTSSIIITTCIPSSGHSRHRYPALIPPLAVITTCIVLYMNGILKC 240

Qy 241 DRKPDRTNSN 250
 Db 241 DRKPDRTNSN 250

RESULT 8
 AAR34371
 ID AAR34371 standard; Protein; 250 AA.
 AC AAR34371;
 XX
 XX
 DT 10-AUG-1993 (first entry)
 XX

XX PF 05-DEC-2000; 2000US-0730465.
 XX PR 07-OCT-1991; 91US-0770969.
 XX PR 02-APR-1992; 92US-0862022.
 XX PR 06-OCT-1992; 92WO-US08755.
 XX PR 06-JUN-1995; 95US-0466465.
 XX PA (BIOJ) BIOGEN INC.
 XX PI Wallner BP, Cooper KD;
 XX WPI: 2002-179118/23.
 XX DR N-PSDB; ABK15717.
 XX
 XX Treating conditions characterised by increased T cell activation and
 XX abnormal antigen presentation in the dermis and epidermis, e.g. atopic
 XX dermatitis, vitiligo and UV damage, by administering an inhibitor of
 XX the CD2/LFA-3 interaction -
 XX
 XX Claim 22; Page 14-15; 32pp; English.
 XX
 XX This invention relates to a novel method for preventing or treating skin
 XX conditions characterised by increased T cell activation and abnormal
 XX antigen presentation in the dermis and epidermis. The method comprises
 XX inhibition of the CD2/lymphocyte function associated antigen-3 (LFA-3)
 XX interaction by administering to a mammal, including a human, an
 XX inhibitor of the CD2/LFA-3 interaction. These inhibitors may be anti-
 XX LFA-3 antibody homologues, anti-CD2 antibody homologues, soluble LFA-3
 XX peptides, soluble CD-2 polypeptides, small molecules such as
 XX carbohydrates, LFA-3 and CD2 mimetic agents. This method is useful to
 XX treat skin disorders such as cell lymphoma, mycosis fungoides, allergic
 XX and irritant contact dermatitis, lichen planus, alopecia areata,
 XX pyoderma gangrenosum, vitiligo, ocular cicatricial pemphigoid,
 XX psoriasis, UV damage and urticaria. The present sequence represents the
 XX mammalian LFA-3 protein sequence used to create the inhibitors of
 XX CD-2/LFA interaction of the invention.
 XX
 XX Sequence 250 AA;
 XX
 XX Query Match 100.0%; Score 1326; DB 23; Length 250;
 XX Best Local Similarity 100.0%; Pred. No. 7.2e-118;
 XX Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVAGSDAGRALGVLSVWCLLHCFGIFSCFSQQIYGVYGNVTFHVPNSVPLKEVLWKQK 60
 Db 1 MVAGSDAGRALGVLSVWCLLHCFGIFSCFSQQIYGVYGNVTFHVPNSVPLKEVLWKQK 60

Qy 61 DKVAELENSEFRAFSSFKNRVYLDTVSGSLTYNLTSSDEDEYEMESPNITDTMKFFLYV 120
 Db 61 DKVAELENSEFRAFSSFKNRVYLDTVSGSLTYNLTSSDEDEYEMESPNITDTMKFFLYV 120

Qy 121 LESLPSTLTCAITNGSIEVQCMPEHNSHRLIMYSWDCPMEQCKRNSTSIYFKMEND 180
 Db 121 LESLPSTLTCAITNGSIEVQCMPEHNSHRLIMYSWDCPMEQCKRNSTSIYFKMEND 180

Qy 181 LPQKIOCTLSNPLNTSSIIITTCIPSSGHSRHRYPALIPPLAVITTCIVLYMNGILKC 240
 Db 181 LPQKIOCTLSNPLNTSSIIITTCIPSSGHSRHRYPALIPPLAVITTCIVLYMNGILKC 240

Qy 241 DRKPDRTNSN 250
 Db 241 DRKPDRTNSN 250

RESULT 7
 AAU76225
 ID AAU76225 standard; Protein; 250 AA.
 AC AAU76225;
 XX
 XX
 DT 08-MAY-2002 (first entry)
 XX
 XX Lymphocyte function-associated antigen-3 (LFA-3) protein.
 XX
 XX LFA-3; lymphocyte function-associated antigen; CD2; dermatological;
 XX cyrostatic; anti-allergic; anti-inflammatory; ophthalmic; dermis;
 XX epidermis; ski disorder; lymphoma; mycosis fungoides; lichen planus;
 XX allergic dermatitis; irritant contact dermatitis; alopecia areata;
 XX pyoderma gangrenosum; vitiligo; ocular cicatricial pemphigoid;
 XX psoriasis; UV damage; urticaria.
 XX
 XX Mammalia.
 XX
 XX Key Location/Qualifiers
 XX Peptide /note= "Signal peptide"
 XX Region 1..28
 XX 28..108
 XX /note= "Peptide inhibitor of CD2/LFA3
 XX interaction. This sequence is specifically
 XX claimed in claim 22 of the specification"
 XX 28..120
 XX /note= "Peptide inhibitor of CD2/LFA3
 XX interaction. This sequence is specifically
 XX claimed in claim 22 of the specification"
 XX 28..215
 XX /note= "Extracellular domain"
 XX 48..108
 XX /note= "Peptide inhibitor of CD2/LFA3
 XX interaction. This sequence is specifically
 XX claimed in claim 22 of the specification"
 XX 39..250
 XX /note= "Mature LFA-3 protein"
 XX 78..93
 XX /note= "Peptide inhibitor of CD2/LFA3
 XX interaction. This sequence is specifically
 XX claimed in claim 22 of the specification"

US2002009449-A1.
 24-JAN-2002.

```
DE Sequence encoded by a naturally occurring human transmembrane
DE LFA-3 DNA.
XX
XX CD2; LFA-3; transmembrane; phosphatidylinositol; T-cell.
XX
XX Homo sapiens.
XX
XX W09306866-A.
XX
XX 15-APR-1993.
XX
XX 06-OCT-1992; 92WO-US08755.
XX
XX 07-OCT-1991; 91US-0770969.
XX
XX 02-APR-1992; 92US-0862022.
XX
XX (BIOJ ) BIOGEN INC.
XX
XX Cooper KD, Wallner BP;
XX
XX WPI; 1993-134148/16.
XX
XX N-PSDB; AAQ40153.
XX
XX Treatment of antigen-linked skin conditions - using inhibitor of
XX CD2-LFA-3 interactions for treating e.g. psoriasis UV damage, and
XX dermatitis
XX
XX Claim 13; Pages 43-44; 60pp; English.
XX
XX LFA-3 is found on the surface of a wide variety of cells, including
XX human erythrocytes. Two natural forms of LFA-3 have been identified.
XX One form of LFA-3 ("transmembrane LFA-3") is anchored in the cell
XX membrane by a transmembrane hydrophobic domain. cDNA encoding this
XX form of LFA-3 has been cloned and sequenced (see, e.g., Wallner et
XX al., (1987)). Another form of LFA-3 is anchored to the cell
XX membrane via a covalent linkage to phosphatidylinositol ("PI") -
XX containing glycolipid. This latter form has been designated "PI-
XX linked LFA-3", and cDNA encoding this form of LFA-3 has also been
XX cloned and sequenced (Wallner et al., PCT publn. WO 90/02181).
XX The human CD2 (P11) molecule is a 50 kD surface glycoprotein
XX expressed on >95% of thymocytes and virtually all peripheral T
XX lymphocytes. The sequence of a human CD2 gene has been reported
XX (Seed and Aruffo (1987)). CD2 cDNA clones predict a cleaved signal
XX peptide of 24 amino acid residues, an extracellular segment of 185
XX residues, a transmembrane domain of 25 residues and a cytoplasmic
XX region of 117 residues.
XX
XX Sequence 250 AA;
XX
XX Query Match 99.0%; Score 1313; DB 14; Length 250;
XX Best Local Similarity 99.2%; Pred. No. 1.2e-116;
XX Matches 248; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 WVGSDAGRALGVLSVCLLHCFGFTSCFSQQIYGVVGNVTFHVPSPNPLKEVLWKQK 60
XX DB 1 WVGSDAGRALGVLSVCLLHCFGFTSCFSQQIYGVVGNVTFHVPSPNPLKEVLWKQK 60
XX
XX QY 61 DKVAELENSEFRAPSFKNRVYLDTVSGSLTIYNTLTSSDEYEMESPNITDTMKFFLYV 120
XX DB 61 DKVAELENSEFRAPSFKNRVYLDTVSGSLTIYNTLTSSDEYEMESPNITDTMKFFLYV 120
XX
XX QY 121 LESLPSPTLTALTNGSIEVQCMIPHYNSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
XX DB 121 LESLPSPTLTALTNGSIEVQCMIPHYNSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
XX
XX QY 181 LPQKIQCTLSNPLFNNTSSIIITTCIPSSGHSRHRALIPPLAVITTCIVLYMNGILKC 240
XX DB 181 LPQKIQCTLSNPLFNNTSSIIITTCIPSSGHSRHRALIPPLAVITTCIVLYMNGILKC 240
XX
XX QY 241 DRKPDRTNSN 250
XX DB 241 HTKPDRTNSN 250
XX
XX RESULT 10
```

```
RESULT 9
AAR34221
ID AAR34221 standard; Protein; 250 AA.
XX
XX AAR34221;
XX
XX 09-AUG-1993 (first entry)
XX
XX Amino acid sequence of transmembrane LFA-3.
XX
XX LFA-3; CD2 binding protein; graft; xenograft; allograft.
XX
XX Synthetic.
XX
XX W09306852-A.
XX
XX 15-APR-1993.
XX
XX 06-OCT-1992; 92WO-US08754.
XX
XX 07-OCT-1991; 91US-0772705.
XX
XX 12-MAR-1992; 92US-0850706.
XX
XX (BIOJ ) BIOGEN INC.
XX
XX Benjamin CD, Wallner BP;
XX
XX WPI; 1993-134134/16.
XX
XX N-PSDB; AAQ40420.
XX
XX Improving tolerance of transplanted allo:graft or xenograft
XX tissue - using LFA-3 or CD-2 binding protein esp. for humans
XX given cardiac or renal grafts
XX
XX Disclosure; Page 47-48; 68pp; English.
XX
XX A bacteriophage comprising (AAQ40420) is deposited under ATCC 75107.
XX Soluble LFA-3 polypeptides may be derived from AAI-AA187. Preferred
XX soluble LFA-3 polypeptides include AAI-AA92, AAI-AA80, AA50-AA65 and
XX AA20-AA80. Soluble LFA-3 polypeptides may also be derived from the
XX PI-linked form of LFA-3 (AAQ40421) (PI = phosphatidylinositol). A
XX vector comprising a DNA sequence encoding PI-linked LFA-3 is
XX deposited under ATCC 68788.
XX
XX Sequence 250 AA;
XX
XX Query Match 98.6%; Score 1307; DB 14; Length 250;
XX Best Local Similarity 98.8%; Pred. No. 4.6e-116;
XX Matches 247; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 WVGSDAGRALGVLSVCLLHCFGFTSCFSQQIYGVVGNVTFHVPSPNPLKEVLWKQK 60
XX DB 1 WVGSDAGRALGVLSVCLLHCFGFTSCFSQQIYGVVGNVTFHVPSPNPLKEVLWKQK 60
XX
XX QY 61 DKVAELENSEFRAPSFKNRVYLDTVSGSLTIYNTLTSSDEYEMESPNITDTMKFFLYV 120
XX DB 61 DKVAELENSEFRAPSFKNRVYLDTVSGSLTIYNTLTSSDEYEMESPNITDTMKFFLYV 120
XX
XX QY 121 LESLPSPTLTALTNGSIEVQCMIPHYNSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
XX DB 121 LESLPSPTLTALTNGSIEVQCMIPHYNSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
XX
XX QY 181 LPQKIQCTLSNPLFNNTSSIIITTCIPSSGHSRHRALIPPLAVITTCIVLYMNGILKC 240
XX DB 181 LPQKIQCTLSNPLFNNTSSIIITTCIPSSGHSRHRALIPPLAVITTCIVLYMNGILKC 240
XX
XX QY 241 DRKPDRTNSN 250
XX DB 241 DRKPDRTNSN 250
XX
XX RESULT 10
```

```

AA028365
ID   AAR28365 standard; Protein; 250 AA.
XX
AC   AAR28365;
XX
DT   21-APR-1993 (first entry)
XX
DE   Human LFA-3 protein.
XX
KW   LFA-3-like; protein; D2; region; LFA-3-delta-D2; LFA-3; receptor;
KW   cell adhesion; immunoglobulin; superfamily; CD2; antigen; D1;
KW   immunoglobulin-like; domain; transmembrane region; TM region;
KW   cytoplasm region; C region; glycosyl phosphatidylinositol; cysteine;
KW   disulphide bond; lymphocyte; function-associated; antigen 3.
XX
OS   Homo sapiens.
XX
FH   Key Location/Qualifiers
FT   Peptide 1..28 "Signal peptide"
FT   Protein 29..250
FT   Region 122..209 /note= "Mature protein"
FT   Region /label D2_region
FT   Region 215..239 /label= TM_region
XX
PN   EP517174-A.
XX
PD   09-DEC-1992.
XX
PF   03-JUN-1992; 92EP-0109340.
XX
PR   06-JUN-1991; 91JP-0134789.
PR   24-JUN-1991; 91JP-0151792.
PR   26-JUN-1991; 91JP-0134486.
PR   02-JUL-1991; 91JP-0161377.
PR   15-JUL-1991; 91JP-0173765.
PR   29-NOV-1991; 91JP-0315709.
XX
PA   (KANF ) KANEGAFUCHI KAGAKU KOGYO KK.
XX
PI   Edamura T, Fukuchi T, Kakutani T, Niwa H, Ohara T;
PI   Osakada F, Yamashita K;
XX
DR   WPI; 1992-408655/50.
DR   N-PSDB; AAQ31676.
XX
PT   Sheep LFA-3 protein and its derivs. - have high affinity for
PT   T-cells, useful for treating T-cell tumours and leukaemia and
PT   also in diagnosis
XX
PS   Disclosure; Page 32-33; 59pp; English.
XX
CC   This sequence represents human lymphocyte function-associated antigen
CC   3 (LFA-3). LFA-3 is a cell adhesion protein belonging to the immuno-
CC   globulin superfamily which acts as a receptor for CD2 antigen. LFA-3
CC   has the structure immuno-globulin-like domain 1 (D1 region), immuno-
CC   globulin-like domain 2 (D2 region), transmembrane region (TM region)
CC   and cytoplasm region (C region). A related CD2 antigen receptor has
CC   the structure D1 region and is bound to the membrane through
CC   glycosyl phosphatidylinositol. LFA-3 has a plurality of disulphide
CC   bonds due to the occurrence of 6 cysteine residues in the D2 region.
CC   See also AAQ31663-90 and AAR28363-70.
XX
SQ   Sequence 250 AA;

Query Match 97.2%; Score 1289; DB 13; Length 250;
Best Local Similarity 97.6%; Pred. No. 2.4e-114;
Matches 244; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Oy 1 MVAGSDAGRALGVLSVCLLHCFISFSQOIYGVVYGNVTFHVPNSVPLKEVLWKQK 60
|||||
1 120
61 DKVAELENSEFRAPSSFKNRVYLDVTSGSLTYNLTSSDEDEYEMESPNTDTMKFFLYV 120
|||||
61 120
61 DKVAELENSEFRAPSSFKNRVYSDVTSGSLTYNLTSSDEDEYEMESPNTDTMKFFLYV 120
|||||
121 LESLPSPTLTCAITNGSIEVQCMIEPHYNSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
|||||
121 180
121 VESLPSPTLTCAITNGSIEVQCMIEPHYNSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
|||||
181 LPQKIQTCLNPLFNTSSIIITTCIPSSGSHRHYALIPILAVITTCIVLYMNGILKC 240
|||||
181 240
181 LPQKIQTCLNPLFNTSSIIITTCIPSSGSHRQYALIPILAVNTTCIVLYMNGILKC 240
|||||
241 DRKPDRTNSN 250
|||||
241 250
DRKPDRTNSN 250

RESULT 11
AAR07604
ID   AAR07604 standard; protein; 240 AA.
XX
AC   AAR07604;
XX
DT   06-FEB-1991 (first entry)
XX
DE   Lymphocyte function-associated antigen (LFA-3) with
DE   phosphatidylinositol (PI) linkage signaling sequence.
XX
KW   Plasma membrane binding affinity; micelle.
XX
FH   Key Location/Qualifiers
FT   Protein 1..189
FT   Peptide /label=LFA-3 Protein
FT   Peptide 190..240
FT   Peptide /label=PI signal sequence
XX
PN   WO9012099-A.
XX
PD   18-OCT-1990.
XX
PF   05-APR-1990; 90WO-US01859.
XX
PR   10-APR-1989; 89US-0335688.
XX
PA   (BIOG-) BIOGEN INC.
XX
PI   Wallner BP;
XX
DR   WPI; 1990-334849/44.
DR   N-PSDB; AAQ06403.
XX
PT   Phosphatidyl-inositol linkage signalling DNA sequence - derived
PT   from lymphocyte function-associated antigen 3, used for prodn. of
PT   chimeric proteins
XX
PS   Disclosure; Fig 2; 53pp; English.
XX
CC   The signal sequence is attached downstream of the LFA-3 sequence
CC   encoding a secreted protein, which will then produce proteins
CC   covalently anchored to the cell surface in which they are produced.
CC   This can give rise to plasma membrane binding, enhanced
CC   purifiability, micelle formation etc. especially useful in the
CC   production of chimeric targeted drugs, to produce micellar or
CC   liposomal delivery systems or in enhanced purification and screening
CC   of cells, proteins or DNA libraries.
XX
SQ   Sequence 240 AA;

Query Match 94.3%; Score 1250; DB 11; Length 240;
Best Local Similarity 99.6%; Pred. No. 1.2e-110;
Matches 236; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 MWAGSDAGRALGVLSVCLLHCFGFTSCFQQIYGVVYGNVTFHVPSPNPLKEVLWKQK 60
 Db 1 MWAGSDAGRALGVLSVCLLHCFGFTSCFQQIYGVVYGNVTFHVPSPNPLKEVLWKQK 60
 QY 61 DKVAELENSEPRAPSFKNRVYLDTVSGSLTIYNLTSSDEDEYEMESPNITDTMKFFLYV 120
 Db 61 DKVAELENSEPRAPSFKNRVYLDTVSGSLTIYNLTSSDEDEYEMESPNITDTMKFFLYV 120
 QY 121 LESLPSPTLTCAITNGSIEVQCMIPHYNSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
 Db 121 LESLPSPTLTCAITNGSIEVQCMIPHYNSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
 QY 181 LPQKIOCTLSNPLFNTSSIIILTTCTIPSSGHSRHRALIPILAVITTCIVLYMNGI 237
 Db 181 LPQKIOCTLSNPLFNTSSIIILTTCTIPSSGHSRHRALIPILAVITTCIVLYMNGM 237

RESULT 12
 AAR27162
 ID AAR27162 standard; Protein; 240 AA.
 XX
 AC AAR27162;
 DT 20-MAY-1998 (first entry)
 XX
 DE PI-linked LFA-3 used to make PIM3 deletion mutant.
 XX
 KW phosphatidylinositol anchored lymphocyte associated antigen-3;
 KW T-lymphocyte accessory molecule; CD2 binding site; immunomodulator;
 KW ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..28
 FT protein /label= signal
 FT 29..240 /label= PI-LFA-3
 FT region 118..188 /label= PIM3
 FT /label= "deleted in mutant"
 XX
 PN EP503648-A.
 XX
 PD 16-SEP-1992.
 XX
 PF 12-MAR-1992; 92EP-0104320.
 XX
 PR 12-MAR-1991; 91US-0667971.
 PR 07-OCT-1991; 91US-0770967.
 XX
 PA (BIOJ) BIOGEN INC.
 XX
 PI Miller GT, Rosa MD, Wallner BP;
 XX
 DR WPI; 1992-309760/38.
 DR N-PSDB; AAQ28683.
 XX
 PT CD2-binding domain of lymphocyte function associated antigen-3
 PT and DNA -- for diagnosing and treating inflammation and
 PT auto-immune diseases, e.g. systemic lupus erythematosus and
 PT rheumatoid arthritis
 XX
 PS Example 9-10; Fig 9; 85pp; English.
 XX
 CC A deletion mutant of PI-LFA-3 was prepared which provided a
 CC PI-linked surface polypeptide on CHO cells having the N-terminal 89
 CC amino acids of native LFA-3. The mutant was generated by deletion
 CC of the region labelled "PIM3" (see features). The PIM3 deletion
 CC mutant form of LFA-3 expressed in a cell line designated PIM3.25.2
 CC exhibited epitopes involved with CD2/LFA-3 complex formation. The
 CC 89 amino acid N-terminal region of the mature LFA-3 molecule

CC contains the conformational requirements necessary for CD2/LFA-3
 CC complex formation.
 XX
 SQ Sequence 240 AA;
 Query Match 94.3%; Score 1250; DB 13; Length 240;
 Best Local Similarity 99.6%; Pred. No. 1.2e-110;
 Matches 236; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MWAGSDAGRALGVLSVCLLHCFGFTSCFQQIYGVVYGNVTFHVPSPNPLKEVLWKQK 60
 Db 1 MWAGSDAGRALGVLSVCLLHCFGFTSCFQQIYGVVYGNVTFHVPSPNPLKEVLWKQK 60
 QY 61 DKVAELENSEPRAPSFKNRVYLDTVSGSLTIYNLTSSDEDEYEMESPNITDTMKFFLYV 120
 Db 61 DKVAELENSEPRAPSFKNRVYLDTVSGSLTIYNLTSSDEDEYEMESPNITDTMKFFLYV 120
 QY 121 LESLPSPTLTCAITNGSIEVQCMIPHYNSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
 Db 121 LESLPSPTLTCAITNGSIEVQCMIPHYNSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
 QY 181 LPQKIOCTLSNPLFNTSSIIILTTCTIPSSGHSRHRALIPILAVITTCIVLYMNGI 237
 Db 181 LPQKIOCTLSNPLFNTSSIIILTTCTIPSSGHSRHRALIPILAVITTCIVLYMNGM 237

RESULT 13
 AAY83134
 ID AAY83134 standard; Protein; 240 AA.
 XX
 AC AAY83134;
 DT 24-JUL-2000 (first entry)
 XX
 DE PI-linked human transmembrane LFA-3.
 XX
 KW LFA3; CD2; cell signalling; modulation; lymphocyte; T cell;
 KW memory effector T lymphocyte; psoriatic arthritis;
 KW rheumatoid arthritis; multiple sclerosis; atopic dermatitis;
 KW uveitis; inflammatory bowel disease; Crohn's disease;
 KW ulcerative colitis; cutaneous T cell lymphoma; inhibition; treatment;
 KW therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO200012113-A2.
 XX
 PD 09-MAR-2000.
 XX
 PF 31-AUG-1999; 99WO-US20026.
 PR 31-AUG-1998; 98US-0098456.
 XX
 PA (BIOJ) BIOGEN INC.
 XX
 PI Magilavy D;
 XX
 DR WPI; 2000-282928/24.
 DR N-PSDB; AAZ93399.
 XX
 PT Selective modulation of memory effector T lymphocytes by administration
 PT of a CD2 binding agent which inhibits the CD2/LFA-3 interaction useful
 PT for treating conditions such as inflammatory bowel diseases, psoriatic
 PT arthritis
 XX
 PS Disclosure; Page 69-70; 76pp; English.
 XX
 CC Modulation of LFA3/CD2 interaction by administration of a CD2 binding
 CC agent inhibits CD2 signalling and T cell proliferation and activation
 CC and more particularly modulates the number and/or distribution of
 CC memory effector T lymphocytes. The method can be used for treating a
 CC condition in a subject where the condition is characterized by memory
 CC effector T lymphocytes playing a role in the pathogenesis of the


```
XX 07-OCT-1991; 91US-0770969.
PR 02-APR-1992; 92US-0862022.
PR 06-OCT-1992; 92WO-US08755.
PR 06-JUN-1995; 95US-0466465.
XX
XX (BIOJ ) BIOGEN INC.
XX
XX Wallner BP, Cooper KD;
XX
XX WPI: 2002-179118/23.
XX N-PSDB; ABR15718.
XX
XX Treating conditions characterised by increased T cell activation and
PT abnormal antigen presentation in the dermis and epidermis, e.g. atopic
PT dermatitis, vitiligo and UV damage, by administering an inhibitor of
PT the CD2/LFA-3 interaction -
XX
XX Example 3; Page 16-17; 32pp; English.
XX
XX This invention relates to a novel method for preventing or treating skin
CC conditions characterised by increased T cell activation and abnormal
CC antigen presentation in the dermis and epidermis. The method comprises
CC inhibition of the CD2/Lymphocyte function associated antigen-3 (LFA-3)
CC interaction by administering to a mammal, including a human, an
CC inhibitor of the CD2/LFA-3 interaction. These inhibitors may be anti-
CC LFA-3 antibody homologues, anti-CD2 antibody homologues, soluble LFA-3
CC peptides, soluble CD-2 polypeptides, small molecules such as
CC carbohydrates, LFA-3 and CD2 mimetic agents. This method is useful to
CC treat skin disorders such as cell lymphoma, mycosis fungoides, allergic
CC and irritant contact dermatitis, lichen planus, alopecia areata,
CC pyoderma gangrenosum, vitiligo, ocular cicatricial pemphigoid,
CC psoriasis, UV damage and urticaria. The present sequence represents the
CC mammalian PI-linked LFA-3 protein sequence used to create the inhibitors
CC of CD-2/LFA interaction of the invention. This protein is anchored to
CC the cell membrane by covalent linkage to a phosphatidylinositol (PI)
CC linkage.
XX
XX SQ Sequence 240 AA;
XX
XX Query Match 94.3%; Score 1250; DB 23; Length 240;
XX Best Local Similarity 99.6%; Pred. No. 1.2e-110;
XX Matches 236; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
Qy 1 MVAGSDAGRALGVLSVCLLHCFGFTSCFSQQIYGVYGVNVTFFHVPSPNPLKEVLWKKQK 60
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1 MVAGSDAGRALGVLSVCLLHCFGFTSCFSQQIYGVYGVNVTFFHVPSPNPLKEVLWKKQK 60
Qy 61 DKVAELENSEFRASFKNRVYLDTVSGSLTYNLTSSDEDEYEMESPNITDTMKFFLYV 120
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61 DKVAELENSEFRASFKNRVYLDTVSGSLTYNLTSSDEDEYEMESPNITDTMKFFLYV 120
Qy 121 LESLPSPTTTCALTNGSIEVQCMIPHYNSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 LESLPSPTTTCALTNGSIEVQCMIPHYNSHRGLIMYSWDCPMEQCKRNSTSIYFKMEND 180
Qy 181 LPQKIOCTLNSPLNFTSSILITCIPSSGCHSRHRYALPIPLAVITTCIVLYMNGI 237
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
181 LPQKIOCTLNSPLNFTSSILITCIPSSGCHSRHRYALPIPLAVITTCIVLYMNGM 237
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 08:38:45 ; Search time 22.6683 Seconds
(without alignments)
634.907 Million cell updates/sec

Title: US-09-730-465-8
Perfect score: 1856
Sequence: 1 MVAGSDAGRALGVLSVCL.....MHEALHNHYTKRSLSPGK 347

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1233	66.4	330	GCL_HUMAN	P01857 homo sapien
2	1145.5	61.7	326	GCL_HUMAN	P01859 homo sapien
3	1135	61.2	327	GCL_HUMAN	P01861 homo sapien
4	1126	60.7	290	GCL_HUMAN	P01860 homo sapien
5	926.5	49.9	323	GCL_RABIT	P01870 oryctolagus
6	892	48.1	329	GCL_CAVPO	P01862 cavia porce
7	845.5	45.6	329	GCL_MOUSE	P22436 mus musculu
8	838	45.2	333	GCL_RAT	P20761 rattus norv
9	834.5	45.0	398	GCL_MOUSE	P03987 mus musculu
10	819	44.1	326	GCL_MOUSE	P20759 rattus norv
11	818.5	44.1	324	GCL_MOUSE	P01868 mus musculu
12	813.5	43.8	393	GCL_MOUSE	P01869 mus musculu
13	812	43.8	329	GCL_MOUSE	P20762 rattus norv
14	811.5	43.7	330	GCL_MOUSE	P01863 mus musculu
15	806.5	43.5	399	GCL_MOUSE	P01865 mus musculu
16	804	43.3	335	GCL_MOUSE	P01864 mus musculu
17	787.5	42.4	322	GCL_MOUSE	P20760 rattus norv
18	783.5	42.2	336	GCL_MOUSE	P01866 mus musculu
19	778.5	41.9	405	GCL_MOUSE	P01867 mus musculu
20	630	33.9	250	LEA3_HUMAN	P19256 homo sapien
21	368	19.8	421	EPC_MOUSE	P06336 mus musculu
22	366.5	19.7	429	EPC_MOUSE	P01855 rattus norv
23	360.5	19.4	428	EPC_MOUSE	P01854 homo sapien
24	358.5	19.3	454	MUC_MOUSE	P01871 homo sapien
25	357	19.2	455	MUC_MOUSE	P01872 mus musculu
26	356.5	19.2	391	MUC_MOUSE	P04220 homo sapien
27	351	18.9	458	MUC_MOUSE	P03988 oryctolagus
28	347	18.7	476	MUC_MOUSE	P01873 mus musculu
29	341	18.4	479	MUC_MOUSE	P04221 oryctolagus
30	335	18.0	457	MUC_MOUSE	P20768 suncus murl
31	332	17.9	450	MUC_MOUSE	P01874 canis famul
32	332	17.9	450	MUC_MOUSE	P06337 mesocricetu
33	302.5	16.3	438	HVC2_HETFR	P23085 heterodontu

34 298.5 16.1 438 1 HVCS_HETFR P23087 heterodontu
35 295.5 15.9 299 1 ALC_RABIT P01879 oryctolagus
36 291 15.7 446 1 MUC_CHICK P01875 gallus gall
37 289.5 15.6 461 1 HVC_MOUSE P23088 heterodontu
38 282.5 15.2 393 1 HVC3_HETFR P23086 heterodontu
39 277.5 15.0 370 1 HVC1_HETFR P23084 heterodontu
40 272 14.7 353 1 ALC1_HUMAN P01876 homo sapien
41 271.5 14.6 340 1 ALC2_HUMAN P01877 homo sapien
42 268 14.4 353 1 ALC1_GORGO P20758 gorilla gor
43 257.5 13.9 481 1 MUC_MOUSE P23735 ictalurus p
44 245 13.2 344 1 ALC_MOUSE P01878 mus musculu
45 177.5 9.6 105 1 LAC1_MOUSE P01843 mus musculu

ALIGNMENTS

RESULT 1
GCL_HUMAN
ID GCL_HUMAN STANDARD; PRT; 330 AA.
AC P01857;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig gamma-1 chain C region.
GN IGHG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82274238; PubMed=6287432;
RA Ellison J.W., Berson B.J., Hood L.E.;
RT "The nucleotide sequence of a human immunoglobulin C gamma gene.";
RL Nucleic Acids Res. 10:4071-4079(1982).
RN [2]
RP SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).
RX MEDLINE=71064024; PubMed=5489771;
RA Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,
RA Maxdal M.J., Edelman G.M.;
RT "The covalent structure of a human gamma G-immunoglobulin. VII. Amino acid sequence of heavy-chain cyanogen bromide fragments H1-H4.";
RL Biochemistry 9:3161-3170(1970).
RN [3]
RP SEQUENCE OF 136-329 (EU).
RX MEDLINE=71064025; PubMed=5530842;
RA Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,
RA Edelman G.M.;
RT "The covalent structure of a human gamma G-immunoglobulin. 8. Amino acid sequence of heavy-chain cyanogen bromide fragments H5-H7.";
RL Biochemistry 9:3171-3181(1970).
RN [4]
RP SEQUENCE (MYELOMA PROTEIN NIE).
RX MEDLINE=77070269; PubMed=826475;
RA Ponstingl H., Hilschmann N.;
RT "The rule of antibody structure. The primary structure of a monoclonal IgG1 immunoglobulin (myeloma protein NIE). III. The chymotryptic peptides of the H-chain, alignment of the tryptic peptides and discussion of the complete structure.";
RL Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).
RN [5]
RP SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.
RX MEDLINE=83289131; PubMed=6884994;
RA Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;
RT "Three-dimensional structure determination of antibodies. Primary structure of crystallized monoclonal immunoglobulin IgG1 KOL, I.";
RL Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).
RN [6]
RP DISULFIDE BONDS.
RX MEDLINE=71064027; PubMed=4923144;
RA Gall W.E., Edelman G.M.;
RT "The covalent structure of a human gamma G-immunoglobulin. X.

RN [2]
 RP SEQUENCE OF 88-115 FROM N.A.
 RC TISSUE=Fetal liver;
 RX MEDLINE=83001943; PubMed=6811139;
 RA Takahashi N., Ueda S., Ohta M., Nikaide T., Nakai S., Honjo T.;
 RT "Structure of human immunoglobulin gamma genes: implications for
 RL evolution of a gene family."; Cell 29:671-679(1982).
 RN [3]
 RP SEQUENCE OF 99-177 AND 310-326 FROM N.A.
 RC TISSUE=Fetal liver;
 RX MEDLINE=84235992; PubMed=6329676;
 RA Krawinkel U., Rabbitts T.H.;
 RT "Comparison of the hinge-coding segments in human immunoglobulin gamma
 RL heavy chain genes and the linkage of the gamma 2 and gamma 4 subclass
 genes."; EMBO J. 1:403-407(1982).
 RN [4]
 RP SEQUENCE OF 1-325 (MYELOMA PROTEIN TIL).
 RX MEDLINE=81007873; PubMed=6774012;
 RA Wang A.-C., Tung E., Fudenberg H.H.;
 RT "The primary structure of a human IgG2 heavy chain: genetic,
 RL evolutionary, and functional implications."; J. Immunol. 125:1048-1054(1980).
 RN [5]
 RP SEQUENCE OF 1-85 AND 132-325 (MYELOMA PROTEIN ZIE).
 RX MEDLINE=80001357; PubMed=113060;
 RA Connell G.E., Parr D.M., Hofmann T.;
 RT "The amino acid sequences of the three heavy chain constant region
 RL domains of a human IgG2 myeloma protein."; Can. J. Biochem. 57:758-767(1979).
 RN [6]
 RP SEQUENCE OF 238-275 (ZIE).
 RX MEDLINE=80114419; PubMed=118920;
 RA Hofmann T., Parr D.M.;
 RT "A note of the amino acid sequence of residues 381-391 of human
 RL immunoglobulin gamma chains."; Mol. Immunol. 16:923-925(1979).
 RN [7]
 RP REVISIONS TO 25; 59; 60 AND 264-268 (ZIE).
 RA Hofmann T., Parr D.M.;
 RL Submitted (MAR-1980) to the PIR data bank.
 RN [8]
 RP SEQUENCE OF 1-121 (DOT).
 RX MEDLINE=95255298; PubMed=7737190;
 RA Stoppini M., Bellotti V., Negri A., Merlini G., Garver F., Ferri G.;
 RT "Characterization of the two unique human anti-flavin monoclonal
 RL immunoglobulins."; Eur. J. Biochem. 228:886-893(1995).
 RN [9]
 RP DISULFIDE BONDS.
 RX MEDLINE=72033500; PubMed=4940472;
 RA Milstein C., Frangione B.;
 RT "Disulphide bridges of the heavy chain of human immunoglobulin G2."; Biochem. J. 121:217-225(1971).
 RN [10]
 RP DISULFIDE BONDS.
 RX MEDLINE=69064124; PubMed=5782707;
 RA Frangione B., Milstein C., Pink J.R.L.;
 RT "Structural studies of immunoglobulin G."; Nature 221:145-148(1969).
 RN [11]
 RP This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 RN [12]
 RP EMBL; J00230; AAB59393.1; -
 DR PIR; A02148; G2HU.
 DR HSP; P01857; 1FC1.

DR Genew; HGNC:5526; IGHG2.
 DR MIM; 147110; -
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003597; Ig_C1.
 DR InterPro; IPR003600; Ig_Like.
 DR Pfam; PF00047; Ig; 3.
 DR SMART; SM00410; Ig_Like; 1.
 DR SMART; SM00407; IGC1; 2.
 DR PROSITE; PS00290; IG_MHC; 2.
 KW Immunoglobulin domain; Immunoglobulin C region.
 FT NON_TER 1 1
 FT DOMAIN 1 98 CHL.
 FT DOMAIN 99 110 HINGE.
 FT DOMAIN 111 219 CH2.
 FT DOMAIN 220 326 CH3.
 FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).
 FT DISULFID 27 83 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 103 103 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 140 200 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 246 304
 FT SITE 156 156
 FT MOD_RES 326 326
 FT VARIANT 60 60
 FT CONFLICT 109 109 /FTID=VAR_003889.
 FT C >> S (IN REF. 3).
 SQ SEQUENCE 326 AA; 35884 MW; 8310878C6879C CRC64;
 Query Match 61.7%; Score 1145.5; DB 1; Length 326;
 Best Local Similarity 91.0%; Pred. No. 9.5e-80;
 Matches 213; Conservative 8; Mismatches 6; Indels 7; Gaps 2;
 Qy 120 VDKT-----HPCPCPAPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEV 173
 Db 94 VDKTVKCCVCPCCAPP-VAGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEV 152
 Qy 174 KFNWYDGVGEVHNAKTPREEQNSTYRVSVLTVLHODWLNKGYCKVSKNKPAPIE 233
 Db 153 QFNWYDGVGEVHNAKTPREEQNSTFRVSVLTVLHODWLNKGYCKVSKNKPAPIE 212
 Qy 234 KTISKAGOPRPQVYTLPPSRDELTKNOVSLTCLVKGYPSDIAVWESNGQPNNYKT 293
 Db 213 KTISKAGOPRPQVYTLPPSRDEMTKNQVSLTCLVKGYPSDIAVWESNGQPNNYKT 272
 Qy 294 TPEVLDSDGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPGK 347
 Db 273 TPEVLDSDGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPGK 326
 RESULT 3
 GC4_HUMAN
 ID GC4_HUMAN STANDARD; PRT; 327 AA.
 AC P01861;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ig gamma-4 chain C region.
 GN IGHG4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN NCBI_TaxID=9606;
 RX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83157104; PubMed=6299662;
 RA Ellison J.W., Buxbaum J.N., Hood L.E.;
 RT "Nucleotide sequence of a human immunoglobulin C gamma 4 gene."; DNA 1:11-18(1981).
 RN [2]
 RP SEQUENCE OF 1-30 AND 81-326.
 RX MEDLINE=70207560; PubMed=4192699;

RA Pink J.R.L., Buttery S.H., de Vries G.M., Milstein C.;
 RT "Human immunoglobulin subclasses. Partial amino acid sequence of the
 RL Biochem. J. 117:33-47(1970).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: K01316; AAB59394.1; ALT_INIT.
 DR PIR: A02150; G4HU.
 DR HSP: P01842; 7FAB.
 DR Genew; HGNC: 5528; IGHG4.
 DR MIM: 147130; -.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003597; Ig_cl.
 DR InterPro: IPR003600; Ig_like.
 DR Pfam: PF00047; Ig; 3.
 DR SMART: SM00410; Ig_like; 1.
 DR SMART: SM00407; IGcl; 2.
 DR PROSITE: PS00290; IG_MHC; 2.
 DR Immunoglobulin domain; Immunoglobulin C region.
 KW NON_TER 1 1
 FT DOMAIN 1 98 CH1.
 FT DOMAIN 99 110 HINGE.
 FT DOMAIN 111 220 CH2.
 FT DOMAIN 221 327 CH3.
 FT DISULFID 14 14
 FT DISULFID 27 83 INTERCHAIN (WITH A LIGHT CHAIN).
 FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 141 201
 FT DISULFID 247 305
 SQ SEQUENCE 327 AA; 35940 MW; 3EDBD811EF208E7A CRC64;
 Query Match 61.2%; Score 1135; DB 1; Length 327;
 Best Local Similarity 93.7%; Pred. No. 6e-79; Mismatches 0; Gaps 0;
 Matches 208; Conservative 8;
 QY 126 CPPCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 185
 DB 106 CPSCPAPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 165
 QY 186 NAKTPREQYNSTYRVSVLTTLVHODWLNKGYCKVSKNKPAPTEKTISKAKGP 245
 DB 166 NAKTPREQYNSTYRVSVLTTLVHODWLNKGYCKVSKNKPAPTEKTISKAKGP 225
 QY 246 PQVYLPSPRDLTKNOVSLTCLVKGFPDIAVWESNGQPNENYKTPPPVLDSDGSFF 305
 DB 226 PQVYLPSPRDLTKNOVSLTCLVKGFPDIAVWESNGQPNENYKTPPPVLDSDGSFF 285
 QY 306 LYSKLTVDKSWQOGNVSFCSVMHEALHNHYTQKSLSLSPK 347
 DB 286 LYSKLTVDKSWQOGNVSFCSVMHEALHNHYTQKSLSLSPK 327
 RESULT 4
 GC3_HUMAN
 ID GC3_HUMAN STANDARD; PRT; 290 AA.
 AC P01860;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ig gamma-3 chain C region (heavy chain disease protein) (HDC).
 GN IGHG3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NX NCBI_TaxID=9606;

RN [1]
 RP SEQUENCE (DISEASE PROTEIN WIS).
 RX MEDLINE=81021548; PubMed=6774747;
 RA Frangione B., Rosenwasser E., Prelli F., Franklin E.C.;
 RT "Primary structure of human gamma 3 immunoglobulin deletion mutant:
 RL gamma 3 heavy-chain disease protein WIS.";
 RN Biochemistry 19:4304-4308(1980).
 CC [2]
 CC REVISIONS TO 12-97 (PROTEIN WIS).
 CC MEDLINE=77118561; PubMed=402363;
 CC Michaelisen T.E., Frangione B., Franklin E.C.;
 CC "Primary structure of the 'hinge' region of human IgG3. Probable
 CC quadruplication of a 15-amino acid residue basic unit.";
 CC J. Biol. Chem. 252:883-889(1977).
 CC [3]
 CC REVISIONS TO 59-289 (PROTEIN WIS/DISEASE PROTEIN ZUC).
 CC MEDLINE=77021516; PubMed=823945;
 CC Wolfenstein-Todel C., Frangione B., Prelli F., Franklin E.C.;
 CC "The amino acid sequence of 'heavy chain disease' protein ZUC.
 CC Structure of the FC fragment of immunoglobulin G3";
 CC Biochem. Biophys. Res. Commun. 71:907-914(1976).
 CC [4]
 CC SEQUENCE FROM N.A. (DISEASE PROTEIN OMM).
 CC MEDLINE=82247835; PubMed=6808505;
 CC Alexander A., Steinmetz M., Barritault D., Frangione B.,
 CC Franklin E.C., Hood L., Buxbaum J.N.;
 CC "Gamma Heavy chain disease in man: cDNA sequence supports partial
 CC gene deletion model";
 CC Proc. Natl. Acad. Sci. U.S.A. 79:3260-3264(1982).
 CC -1- SUBUNIT: DIMER LINKED BY 12 DISULFIDE BONDS; IT HAS AN EXTRA
 CC INTERCHAIN DISULFIDE BOND AT POSITION 7 IN ADDITION TO THE 11
 CC NORMALLY PRESENT IN THE HINGE REGION.
 CC -1- MISCELLANEOUS: THE HEAVY CHAIN DISEASE PROTEIN WIS IS SHOWN.
 CC -1- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 42-76 WAS TAKEN FROM THE
 CC REF. 2.
 CC -1- MISCELLANEOUS: DISEASE PROTEIN WIS IS LACKING MOST OF THE V REGION
 CC AND ALL OF THE CH1 REGION.
 CC -1- MISCELLANEOUS: DISEASE PROTEIN ZUC LACK MOST OF THE V REGION, ALL
 CC OF THE CH1 REGION, AND PART OF THE HINGE COMPARED WITH NORMAL
 CC GAMMA-3 HEAVY CHAINS.
 CC -1- MISCELLANEOUS: DISEASE PROTEIN OMM MAY REPRESENT AN ALLELIC FORM
 CC OR ANOTHER GAMMA CHAIN SUBCLASS.
 CC -1- MISCELLANEOUS: THE HINGE REGION IN GAMMA-3 CHAINS IS ABOUT FOUR
 CC TIMES AS LONG AS IN OTHER GAMMA CHAINS AND CONTAINS THREE
 CC IDENTICAL 15-RESIDUE SEGMENTS PRECEDED BY A SIMILAR 17-RESIDUE
 CC SEGMENT (12-28).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: J00231; AAA52805.1; ALT_SEQ.
 CC PIR: A02149; G3HUMI.
 CC HSP: P01857; 1FC1.
 CC Genew; HGNC: 5527; IGHG3.
 CC MIM: 147120; -.
 CC InterPro: IPR003006; Ig_MHC.
 CC InterPro: IPR003597; Ig_cl.
 CC InterPro: IPR003600; Ig_like.
 CC Pfam: PF00047; Ig; 2.
 CC SMART: SM00410; Ig_like; 1.
 CC SMART: SM00407; IGcl; 1.
 CC PROSITE: PS00290; IG_MHC; 1.
 CC Immunoglobulin domain; Immunoglobulin C region; Glycoprotein; Repeat.
 KW DOMAIN 12 73 HINGE.
 FT DOMAIN 74 183 CH2.
 FT DOMAIN 184 289 CH3.
 FT REPEAT 29 43
 FT REPEAT 44 58

```
FT REPEAT 59 73 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 1 1 N-LINKED (GLNAC...).
FT CARBOHYD 6 6 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 7 7 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 24 24 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 27 27 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 33 33 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 39 39 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 42 42 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 48 48 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 54 54 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 57 57 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 63 63 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 69 69 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 72 72 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT CARBOHYD 140 140 N-LINKED (GLNAC...).
FT MOD_RES 290 290 REMOVED POST-TRANSLATIONALLY.
FT VARIANT 126 127 QV -> ER (IN ZUC).
FT VARIANT 134 134 P -> L (IN OMM).
FT VARIANT 139 139 /FTId=VAR_003890.
FT VARIANT 182 182 F -> Y (IN OMM).
FT VARIANT 227 227 /FTId=VAR_003891.
FT VARIANT 227 227 T -> A (IN OMM).
FT VARIANT 227 227 S -> N (IN OMM).
FT VARIANT 227 227 /FTId=VAR_003893.
FT VARIANT 227 227 /FTId=VAR_003894.
FT VARIANT 227 227 MISSING (IN ZUC).
FT VARIANT 279 279 /FTId=VAR_003895.
FT VARIANT 279 279 F -> Y (IN OMM).
FT VARIANT 279 279 /FTId=VAR_003896.
SQ SEQUENCE 290 AA; 32331 MW; E69CB95705B2R46 CRC64;

Query Match 60.7%; Score 1126; DB 1; Length 290;
Best Local Similarity 90.3%; Pred. No. 2.5e-78;
Matches 205; Conservative 13; Mismatches 9; Indels 0; Gaps 0;

QY 121 DKTHPCPCPAPELLGGPSVFLPPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYD 180
DB 64 DTPPPCPPCPAPELLGGPSVFLPPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYD 123
QY 181 GVEVHNATKPREQYNSTVRYSVLTVLHDLNGLNKEYCKYKSNKALPAPIEKTISKAK 240
DB 124 GVQVHNATKPREQYNSTVRYSVLTVLHDLNGLNKEYCKYKSNKALPAPIEKTISKAK 183
QY 241 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 300
DB 184 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 243
QY 301 DGSFELYSLVDRSRQOQNVFSCSVMEALHNHYTQKSLSLSPGK 347
DB 244 DGSFELYSLVDRSRQOQNVFSCSVMEALHNHYTQKSLSLSPGK 290

RESULT 5
GC_RABIT STANDARD; PRT; 323 AA.
AC F01870;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma chain C region.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84030930; PubMed=6313520;
RA Bernstein K.E.; Alexander C.B.; Mage R.G.;
RT "Nucleotide sequence of a rabbit IgG heavy chain from the recombinant F-I haplotype.";
RL Immunogenetics 18:387-397(1983).
```

```
RN SEQUENCE OF 1-128.
RX MEDLINE=76135469; PubMed=1243651;
RA Pratt D.M.; Mole L.E.;
RT "Sequence studies on the constant region of the Fd sections of rabbit immunoglobulin G of different allotype.";
RL Biochem. J. 151:337-349(1975).
[3]
RP SEQUENCE OF 88-266 FROM N.A.
RX MEDLINE=83299917; PubMed=6193512;
RA Martens C.L.; Moore K.W.; Steinmetz M.; Hood L.; Knight K.L.;
RT "Heavy chain genes of rabbit IgG: isolation of a cDNA encoding gamma heavy chain and identification of two genomic C gamma genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:6018-6022(1982).
[4]
RP SEQUENCE OF 132-161.
RX MEDLINE=70110015; PubMed=5461106;
RA Fruchter R.G.; Jackson S.A.; Mole L.E.; Porter R.R.;
RT "Sequence studies of the Fd section of the heavy chain of rabbit immunoglobulin G.";
RL Biochem. J. 116:249-259(1970).
[5]
RP SEQUENCE OF 129-131 AND 155-322.
RA Hill R.L.; Lebovitz H.E.; Fellows R.E. Jr.; Delaney R.;
RL (In) Killander J. (eds.);
RL Gamma globulins, Nobel symp. 3, pp.109-127, Almqvist and Wiksell, Stockholm (1967).
CC -1- MISCELLANEOUS: REF.1 SEQUENCE HAS THE D12 ALLOTYPIC MARKER, 104-THR, AND THE E14 MARKER, 185-THR. REF.3 HAS THE D11 AND E15 MARKERS AND REF.5 THE E15 MARKER.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC
CC EMBL; M16426; AAA31289.1; -.
CC PTR; A02161; GHRB.
CC DR HSP; P01857; 1FC1.
CC DR InterPro; IPR003006; Ig_MHC.
CC DR InterPro; IPR003597; Ig_C1.
CC DR Pfam; PF00047; Ig_2.
CC DR SMART; SM00407; IgC1; 2.
CC DR PROSITE; PS00290; IG_MHC; 1.
CC KW Immunoglobulin domain; Immunoglobulin C region.
CC NON_TER 1
CC FT VARIANT 104 104 T -> M (IN D11 MARKER).
CC FT VARIANT 185 185 T -> A (IN E15 MARKER).
CC FT CONFLICT 48 48 N -> E (IN REF. 2).
CC FT CONFLICT 71 71 V -> VPV (IN REF. 2).
CC FT CONFLICT 144 144 Q -> E (IN REF. 3 AND 4).
CC FT CONFLICT 173 173 N -> D (IN REF. 5).
CC FT CONFLICT 187 187 Q -> E (IN REF. 3 AND 5).
CC FT CONFLICT 201 201 N -> D (IN REF. 5).
CC FT CONFLICT 218 218 Q -> E (IN REF. 5).
CC FT CONFLICT 233 233 E -> Q (IN REF. 5).
CC FT CONFLICT 246 246 N -> D (IN REF. 5).
CC FT CONFLICT 256 256 E -> G (IN REF. 5).
CC FT CONFLICT 260 260 N -> D (IN REF. 5).
CC FT CONFLICT 266 266 N -> D (IN REF. 5).
CC FT CONFLICT 280 280 Y -> W (IN REF. 5).
CC FT CONFLICT 284 284 N -> S (IN REF. 5).
CC SQ SEQUENCE 323 AA; 35404 MW; 69E8AA118D579A8B CRC64;

Query Match 49.9%; Score 926.5; DB 1; Length 323;
Best Local Similarity 63.3%; Pred. No. 3.7e-63;
Matches 178; Conservative 38; Mismatches 52; Indels 13; Gaps 4;

QY 72 RAFFSKNRVYLDPTVSGSLTIYNLTSSDEDEYEMESPNIITDMKFLYVDKT---HTC-- 126
RL | | | : | : | : : : | | : | : | | |
```

```

Db 51 RTPPSVROSSGLYSLS---SYVSVTSSQPVTCNVHAHPATNTK-----VDKTVAPSTCSK 102
Qy 127 PPCPAPELLGGPSVFLPPPKDITLMSRTPEVTCVVVDVSHEDPEVKENWYVDGVEVHN 186
Db 103 PTCPELLGGPSVFLPPPKDITLMSRTPEVTCVVVDVSHEDPEVKENWYVDGVEVHN 162
Qy 187 AKTKPREQYNTSVRVSVTLVHLDWLNKGYCKVSKNKPAPIEKTIKAKGQPREP 246
Db 163 ARPLREQNFNTSVRVSTLPIHODWLRGKFKCKVHNKALPAPIEKTIKAKGQPLEP 222
Qy 247 QVYTLPPSDELTKNOVSLCLVKGFPSDIAVWESNGPENNKTTPPVLDSGSRFL 306
Db 223 KYVTGPPREELSSRSVLTCHMGYPSPDISVEWENKGAEDNYKTTPVLDSDGSYFL 282
Qy 307 YSKLTVDKSRQOQNVFSCSVHMHYKLNHYTKSLSPGK 347
Db 283 YNKLVSPTSEMQRGDVTCSVMHEALHNHYTKSLSPGK 323

RESULT 6
GC2_CAVPO STANDARD; PRT; 329 AA.
AC P01862;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE Ig gamma-2 chain C region.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
NCBI_TaxID=10141;
RN [1]
RP Trischmann T.M.;
RL Submitted (APR-1975) to the PIR data bank.
[2]
SEQUENCE OF 1-3.
RX MEDLINE=71058471; PubMed=5538606;
RT "Structure of heavy chain from strain 13 guinea pig
immunoglobulin-G(2). 3. Amino acid sequence of the region around the
half-cysteine joining heavy and light chains."
RL Biochemistry 10:18-25(1971).
[3]
RX MEDLINE=71058486; PubMed=5538616;
RT "Structure of heavy chain from strain 13 guinea pig
immunoglobulin-G(2). II. Amino acid sequence of the carboxyl-terminal
and hinge region cyanogen bromide fragments."
RL Biochemistry 10:9-17(1971).
[4]
RX MEDLINE=75036072; PubMed=4429665;
RT "Primary structure of the CH2 homology region from guinea pig IgG2
antibodies."
RL Biochemistry 13:4796-4803(1974).
[5]
RX MEDLINE=75036073; PubMed=4609467;
RT "Primary structure of the CH3 homology region from guinea pig IgG2
antibodies."
RL Biochemistry 13:4804-4811(1974).
[6]
RX MEDLINE=71058474; PubMed=4922544;
RT "Interchain disulfide bridges of guinea pig gamma-2-immunoglobulin."
RL Biochemistry 10:26-31(1971).
CC -|- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM POOLED SERUM OF STRAIN
CC 13 INBRED GUINEA PIGS.

```

```

DR PIR; A02151; GZGP.
DR HSP; P01842; 7FAB.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_C1.
DR InterPro: IPR003600; Ig_Like.
DR Pfam: PF00047; Ig_2.
DR SMART: SM00410; Ig_Like; 1.
DR SMART: SM00407; IGG1; 2.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1
FT DISULFID 16 16 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 28 79
FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 142 202
FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .).
FT DISULFID 248 308
SQ SEQUENCE 329 AA; 5D231B7164D1FBA9 CRC64;

Query Match 48.18; Score 892; DB 1; Length 329;
Best Local Similarity 60.6%; Pred. No. 1.5e-60;
Matches 177; Conservative 30; Mismatches 47; Indels 38; Gaps 6;

Qy 87 SGSLT-----IYNLT-----SSDEDEYEMESPNITDITMKFFLYVDKT----- 123
Db 44 SGALTSVGVHTFPAVLQSLGSLTSMVTVPSSQKATCNVAHP--ASSTK---VDKTVPEI 97
Qy 124 -----HTCPCPAPELLGGPSVFLPPPKDITLMSRTPEVTCVVVDVSHEDPEVKEN 176
Db 98 RTPZPBPCCTCPKPPENLGGPSVFLPPPKDITLMSRTPEVTCVVVDVSHEDPEVQFT 157
Qy 177 WYVDGVEVHNKTKPREQYNTSVRVSVTLVHLDWLNKGYCKVSKNKPAPIEKTI 236
Db 158 WFDNKPVGNAETKPRVBYNTTFVESVLPVHQDMLRGKFKCKVHNKALPAPIEKTI 217
Qy 237 SKAGQPREQYNTSVRVSVTLVHLDWLNKGYCKVSKNKPAPIEKTI--ENNYKTT 294
Db 218 SKTKGAPRPDYTLPPSRDELSKSVYVTCILNFFPADIHVEWASNRVPVSEKYNK 277
Qy 295 PPVLDSGSRFLYKLTVDKSRQOQNVFSCSVHMHYKLNHYTKSLSPG 346
Db 278 PRIEDADGSYFLYKLTVDKSRQOQNVFSCSVHMHYKLNHYTKSLSPG 329

RESULT 7
GC3_MOUSE STANDARD; PRT; 329 AA.
AC P22436;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DE Ig gamma-3 chain C region, secreted form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;
RA Weis J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene."
RL EMBO J. 3:2041-2046(1984).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL: J00451; -; NOT_ANNOTATED_CDS.
DR PIR: B02156; G3MSC.
DR HSSP: P01857; 1FC1.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_Cl.
DR InterPro: IPR003600; Ig_Like.
DR Pfam: PF00047; Ig; 3.
DR SMART: SM00410; Ig_Like; 1.
DR SMART: SM00407; IGc1; 2.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
FT SEQUENCE 329 AA; 36228 MW; F45827174182BAD6 CRC64;

Query Match 45.6%; Score 845.5; DB 1; Length 329;
Best Local Similarity 67.0%; Pred. No. 5.1e-57;
Matches 150; Conservative 34; Mismatches 37; Indels 3; Gaps 1;

QY 127 PP---CPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVE 183
   || || :|||||:||||| |||| ||:||||| |||| ||:||||| |||| ||:|||||
DQ 106 PGSSCPPGNILGGPSVFLPPKPKDMLSLTPKVTCTVVDVSEDDPDVHVSFWFDNKE 165

QY 184 VHNATKPREQYNSTYRVSVLVHLQDWLNKGYKCKVSNKALPAPIETKTSKAKGP 243
   || ||:|||||:||||| |||| ||:||||| |||| ||:||||| |||| ||:|||||
DQ 166 VHTAQTPREQYNSTYRVSVLVHLQDWLNKGYKCKVSNKALPAPIETKTSKAKGP 225

QY 244 REPOVYTLPPSDELTKNQVSLTCLVKGFYPSDIAVEHESNGQPENNYKTPPVLDSDGS 303
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 226 QTPQVYIIPPREQSKKVSILCLVTFSEAISEVWERNGELQDYKNTPTPLSDSGT 285

QY 304 FFLYSLKLTVDKSRQOGNVFSCSVNHEALHNHYTKQSLSLSPGK 347
   :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 286 YFLYSLKLTVDTSWLQGEIFCTSVVHEALHNHHTQKLSRSPGK 329

RESULT 8
GCB_RAT
ID GCB_RAT STANDARD; PRT; 333 AA.
AC P20761;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE Ig gamma-2B chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
DR PIR: PS0018; PS0018.
DR HSSP: P01842; 7FAB.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_Cl.
DR InterPro: IPR003600; Ig_Like.
DR Pfam: PF00047; Ig; 3.
DR SMART: SM00410; Ig_Like; 1.
DR SMART: SM00407; IGc1; 2.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 80
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).

```

```

FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 147 207
FT DISULFID 253 311
SQ SEQUENCE 333 AA; 36497 MW; 55F8B64D48D460A6 CRC64;

Query Match 45.2%; Score 838; DB 1; Length 333;
Best Local Similarity 66.4%; Pred. No. 1.9e-56;
Matches 148; Conservative 33; Mismatches 42; Indels 0; Gaps 0;

QY 125 TCPPCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV 184
   || ||:|||||:||||| |||| ||:||||| |||| ||:||||| |||| ||:|||||
DQ 111 TCHKCPVPELGGPSVFLPPKPKDILLISNAKVTCVVVDVSEEDPDVQFSWNNYVEV 170

QY 185 HNATKPREQYNSTYRVSVLVHLQDWLNKGYKCKVSNKALPAPIETKTSKAKGP 244
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 171 HTAQTPREQYNSTYRVSVLVHLQDWLNKGYKCKVSNKALPAPIETKTSKAKGP 230

QY 245 REPOVYTLPPSDELTKNQVSLTCLVKGFYPSDIAVEHESNGQPENNYKTPPVLDSDGSF 304
   :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 231 KQVYVMGPTEQTEQVTSCLTSGFLPDIGVWTSNGHTIEKYNKTEPVMDSDGSF 290

QY 305 FLYSLKLTVDKSRQOGNVFSCSVNHEALHNHYTKQSLSLSPGK 347
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 291 FVSKLVNVSERWDSRAPFVCSVHGLNHHVSEKISRPFGK 333

RESULT 9
GCB_MOUSE
ID GCB_MOUSE STANDARD; PRT; 398 AA.
AC P03987;
DT 23-OCT-1986 (Rel. 02, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-3 chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;
RA Wells J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBO J. 3:2041-2046(1984).
RN [2]
RP SEQUENCE OF 328-398 FROM N.A.
RX MEDLINE=84041483; PubMed=6314258;
RA Komaromy M., Clayton L., Rogers J., Robertson S., Kettman J.,
RA Walli R.;
RT "The structure of the mouse immunoglobulin in gamma 3 membrane gene
segment.";
RL Nucleic Acids Res. 11:6775-6785(1983).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: J00451; AAB59655.1; -.
CC EMBL: V01526; CAA24767.1; ALT_SEQ.
CC PIR: A02155; G3MSM.
CC HSSP: P01857; 1FC1.
CC InterPro: IPR003006; Ig_MHC.
CC InterPro: IPR003597; Ig_Cl.
CC InterPro: IPR003600; Ig_Like.
CC Pfam: PF00047; Ig; 3.
CC SMART: SM00410; IG_Like; 1.

```

DR SMART; SM00407; IGcl; 2.
 DR PROSITE; PS00290; IG_MHC; 1.
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
 KW Transmembrane; Alternative splicing.
 FT NON_TER 1 1
 FT DOMAIN 1 97 CH1.
 FT DOMAIN 98 113 HINGE.
 FT DOMAIN 114 223 CH2.
 FT DOMAIN 224 327 CH3.
 FT TRANSMEM 346 362 POTENTIAL.
 FT DOMAIN 363 398 CYTOPLASMIC (POTENTIAL).
 FT CONFLICT 333 333 E -> G (IN REF. 2).
 FT CONFLICT 342 342 E -> Q (IN REF. 2).
 FT CONFLICT 388 388 P -> F (IN REF. 2).
 SQ SEQUENCE 398 AA; 43929 MW; CF7264B50A41B95 CRC64;

Query Match 45.0%; Score 834.5; DB 1; Length 398;
 Best Local Similarity 66.7%; Pred. No. 4.4e-56;
 Matches 148; Conservative 34; Mismatches 37; Indels 3; Gaps 1;

QY 127 PP---CPAPELLGSPVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVE 183
 DB 106 PPGSCPPGNLTGSPVFIFPPPKKDALMISLTPEKTCVVVDVSEDDPDVHYSWFDNKE 165
 QY 184 VINAKTKPREQYNSTRVYVSLTVLHODWLNKGEYCKVSNKALPAPIETKISKAKGAP 243
 DB 166 VHTANTQPREAQYNSTRVYVSLALPIQHDWNRGKFEKCKYNNKALPAPIERTISKPKGRA 225
 QY 244 REPQVYTPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTPTPPVLDSGGS 303
 DB 226 QTPQVYTIPTPPREQMSKKVSLTCLVTFNFSSEISVERNGELEDQDYKNTPTPLDSGT 285
 QY 304 FFLYSLKLVVDKSRWQGNVFCSCVMHEALHNHYTOKLSLSP 345
 DB 286 YFLYSLKLVVDKSRWQGNVFCSCVMHEALHNHYTOKLSLSP 327

RESULT 10
 GC1_RAT
 ID GC1_RAT STANDARD; PRT; 326 AA.
 AC P20759;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig gamma-1 chain C region.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89232738; PubMed=3149946;
 RA Bruggemann M.;
 RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
 RL Gene 74:473-482(1988).
 DR PIR; PS0017;
 DR HSSP; P01842; 7FAB.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003597; Ig-cl.
 DR Pfam; PF00047; Ig; 3.
 DR SMART; SM00407; IGcl; 2.
 DR PROSITE; PS00290; IG_MHC; 1.
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
 FT NON_TER 1 1
 FT DOMAIN 1 97 CH1.
 FT DOMAIN 98 112 HINGE.
 FT DOMAIN 113 219 CH2.
 FT DOMAIN 220 326 CH3.
 FT DISULFID 27 82
 FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).

FT DISULFID 140 200
 FT DISULFID 246 304
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 326 AA; 35946 MW; 013BAB45EF49B9DA CRC64;

Query Match 44.1%; Score 819; DB 1; Length 326;
 Best Local Similarity 54.8%; Pred. No. 5.1e-55;
 Matches 153; Conservative 45; Mismatches 45; Indels 36; Gaps 5;

QY 92 IYNLTSSDEYEMESPNTDTMKFFLYVDKTTCTCPG-PA-----PELLGG--- 137
 DB 61 LYLTSS-----VTVPSST-----WPSQVTCNVVAHPASSTKVKKIIVRCNGGDK 107
 QY 138 -----PSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVINAK 188
 DB 108 PCICTGSEVSSVFIFPPPKKDVLTITLTPKVTCTVYVVDISQDDPEVHFVDFVDDVEVHTAQ 167
 QY 189 TKPREQYNSTRVYVSLTVLHODWLNKGEYCKVSNKALPAPIETKISKAKGPREPQV 248
 DB 168 TRPEQGFNSTPRSYSELPIHODWLNGRTRCKVTSAAFPSPLEKTSKPEGRTOVPHV 227
 QY 249 YTPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTPTPPVLDSGGSFFLYS 308
 DB 228 YTMSPTEKEMTQNEVSICTVWKGFPDPIYVWQMNGQFQENYKNTPTMTDTCGYFLYS 287
 QY 309 KLTVDKSRWQGNVFCSCVMHEALHNHYTOKLSLSPGK 347
 DB 288 KLVNKKWQGNVFCSCVMHEALHNHYTOKLSLSPGK 326

RESULT 11
 GC1_MOUSE
 ID GC1_MOUSE STANDARD; PRT; 324 AA.
 AC P01868;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Ig gamma-1 chain C region.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=80045036; PubMed=115593;
 RA Honjo T., Obara M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
 RA Takahashi N., Mano Y.;
 RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
 gamma 1 chain gene.";
 RL Cell 18:559-568(1979).
 RN [2]
 RP SEQUENCE OF 76-324 FROM N.A. (MYELOMA PROTEIN MOPC 31C).
 RX MEDLINE=80202559; PubMed=6769752;
 RA Obara T., Yamawaki-Kataoka Y., Takahashi N., Kataoka T., Shimizu A.,
 RA Mano Y., Seidman J.G., Peterlin B.M., Ieder P., Honjo T.;
 RT "Immunoglobulin gamma 1 heavy chain gene: structural gene sequences
 cloned in a bacterial plasmid.";
 RL Gene 9:87-97(1980).
 RN [3]
 RP SEQUENCE OF 70-322 FROM N.A. (MYELOMA PROTEIN MOPC 21).
 RX MEDLINE=80012837; PubMed=113776;
 RA Rogers J., Clarke P., Salser W.;
 RT "Sequence analysis of cloned cDNA encoding part of an immunoglobulin
 heavy chain.";
 RL Nucleic Acids Res. 6:3305-3321(1979).
 RN [4]
 RP SEQUENCE (MYELOMA PROTEIN MOPC 21).
 RX MEDLINE=78242288; PubMed=98524;
 RA Adetugbo K.;
 RT "Evolution of immunoglobulin subclasses. Primary structure of a
 murine myeloma gammal chain.";
 RL J. Biol. Chem. 253:6068-6075(1978).
 RN [5]

```
RP DLSULFIDE BONDS (MOPC 21).
RX MEDLINE=7300889; PubMed=5073237;
RA Svasti J., Milstein C.;
RT "The disulphide bridges of a mouse immunoglobulin G1 protein.";
RL Biochem. J. 126:837-850(1972).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00793; CAA24172.1; -
DR EMBL; V00793; CAA24173.1; -
DR EMBL; V00793; CAA24174.1; -
DR EMBL; V00793; CAA24175.1; -
DR EMBL; V00793; CAA24176.1; -
DR PIR; A02159; GIMS.
DR HSP; P01842; 7FAB.
DR GlycoSuiteDB; P01868; -.
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_c1.
DR Pfam; PF00047; ig; 3.
DR SMART; SM00407; Igcl; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
FT Alternative splicing.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 110 HINGE.
FT DOMAIN 111 217 CH2.
FT DOMAIN 218 324 CH3.
FT DISULFID 27 82 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 138 198 N-LINKED (GLCNAC. .).
FT CARBOHYD 174 174 /FTID=CAR_000055.
FT DISULFID 244 302
FT MOD_RES 324 324 REMOVED POST-TRANSLATIONALLY.
FT CONFLICT 276 276 N -> D (IN REF. 3).
FT CONFLICT 278 278 N -> D (IN REF. 3).
FT SEQUENCE 324 AA; 35704 MW; A338812F3D1F2C93 CRC64;
Query Match 44.1%; Score 818.5; DB 1; Length 324;
Best Local Similarity 54.1%; Pred. No. 5.6e-55;
Matches 151; Conservative 51; Mismatches 48; Indels 29; Gaps 5;
QY 83 LDTVSGSLTYVNLTSDEDEYEWESPNTIDTMKFFL-----YVDKHT-----CpP-- 128
Db 61 LYTLSSTVTV-----PSSPRSEIVTCNVAHPASSPKVDKIVPRDCGKPCI 108
QY 129 CPAPLGGPSVFLPPPKDILMISRPETVCVVVDVSHEDPEVKFNWYDGVGVINAK 188
Db 109 CTVPEV---SSVFIFFPKPKDVTITLTPKVTVCVVVDISKDDPEVQFSFVDDVEVHTAQ 165
QY 189 TKPREQVNSTYRVVSVLTVLHQDMLNGKEYCKVKVSNKALPAPIEKTIISKARGQPRPQV 248
Db 166 TQPREQVNSTYRVVSVLTVLHQDMLNGKEYCKVKVSNKALPAPIEKTIISKARGQPRPQV 225
QY 249 YTLPSRDELTKNQVSLTCLYGFEYPSDIAVESNGQENYKTPPVLDSDGSFFLYS 308
Db 226 YTIPTPKQMAKQKVSUTCMITDFPDDITVEWMQNGQPAENTKQPIPMNTINGSYFVIS 285
QY 309 KLTVDKSRWQGNQVFCSVNHEALHNHYTQKSLSLSPGK 347
Db 286 KLVNQKSNWEAGNFTTCSVLREGLNHNHTKSLSPGK 324
```

```
RESULT 12
GCIM_MOUSE STANDARD; PRT; 393 AA.
ID GCIM_MOUSE
AC P01869;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ig gamma-1 chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=80045036; PubMed=115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
gamma 1 chain gene.";
RL Cell 18:559-568(1979).
RN [2]
RN SEQUENCE OF 323-393 FROM N.A.
RP MEDLINE=82197626; PubMed=6804950;
RA Tyler B.M., Cowman A.F., Gerondakis S.D., Adams J.M., Bernard O.;
RT "mRNA for surface immunoglobulin gamma chains encodes a highly
conserved transmembrane sequence and a 28-residue intracellular
domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2008-2012(1982).
RN [3]
RN SEQUENCE OF 323-366 FROM N.A.
RP MEDLINE=82115295; PubMed=6799207;
RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,
RA Eisenberg D., Wall R.;
RT "Gene segments encoding transmembrane carboxyl termini of
immunoglobulin gamma chains.";
RL Cell 26:19-27(1981).
RN [4]
RN SEQUENCE OF 1-44 FROM N.A.
RP MEDLINE=82222190; PubMed=6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
immunoglobulin gamma chains.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
CC -!- ALTERNATIVE PRODUCTS; CELL LINES PRODUCING IGG CONTAIN TWO MRNA
CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
CC BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
CC IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
CC SEGMENT OF MU CHAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00793; CAA24172.1; -
DR EMBL; V00793; CAA24173.1; -
DR EMBL; V00793; CAA24174.1; -
DR PIR; B02159; GIMS.
DR HSP; P01842; 7FAB.
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_c1.
DR Pfam; PF00047; ig; 3.
DR SMART; SM00407; Igcl; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing; Transmembrane.
FT NON_TER 1
FT 1
```

```

DR DOMAIN 1 97 CHI.
FT DOMAIN 98 110 HINGE.
FT DOMAIN 111 217 CH2.
FT DOMAIN 218 324 CH3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 138 198
FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .).
FT DISULFID 244 302 POTENTIAL.
FT TRANSMEM 340 357 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 358 393
SQ SEQUENCE 393 AA; 43386 MW; 4CC88343B7A1CE27 CRC64;

Query Match 43.8%; Score 813.5; DB 1; Length 393;
Best Local Similarity 54.0%; Pred. No. 1.7e-54;
Matches 150; Conservative 51; Mismatches 48; Indels 29; Gaps 5;

Qy 83 LDTVSGSLTYNLTSSDEYEMESPNTDTMKFFL-----YVDKTHT-----CPP-- 128
Db 61 LYTLSSTVTV-----PSSRPSETVTCNVAHPASSTKVDKKIVPRDCGCKPCI 108

Qy 129 CPAPELLGGPSVFLPPPKPDTLMISRTPEVTCVVVDVSHEDPEVKENYVDGVEVHNAK 188
Db 109 CTYPEV---SSVFIFFPKPKDVLITLTPKVTCTVVDVSKDDPEVQFSWFVDVVEHTAQ 165

Qy 189 TKPREQYNSTYRVVSVLTVHLQDWLNGKEYCKVSKNALPAIEKTIISKAKQPREQV 248
Db 166 TQPREQFNSTERSSELPIMHODWLNGKEFKCRVNSAAFPAPIEKTIISKAKGRKAPQV 225

Qy 249 YTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGOPENNYKTTTPVLDSDGSFFLYS 308
Db 226 YTIFFPKQMAKDVKVSLTCMTIDFFPEDITVEMQWNGQPAENYKNTQPINNTNGSYFVVS 285

Qy 309 KLIVDKSRQOQNVFSCVMHEALHNHYTKSLSLSPG 346
Db 286 KLVQKSNWEAGNTFTCSVLHLEGLHNHHTEKSLSHSPG 323

RESULT 13
GCC_RAT GCC_RAT STANDARD; PRT; 329 AA.
AC P20762;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2C chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88166903; PubMed=3127222;
RA Brueggemann M., Delmastro-Galire P., Waldmann H., Calabi F.;
RT "Sequence of a rat immunoglobulin gamma 2c heavy chain constant
region cDNA: extensive homology to mouse gamma 3.";
RL Eur. J. Immunol. 18:317-319(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X07189; CAA30169.1; .
DR F01842; 7FAB.
DR HSP; F01842; 7FAB.
DR InterPro; IPR003006; Ig_MHC.

```

```

DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; ig; 2.
DR SMART; SM00410; Igcl; 1.
DR SMART; SM00407; Igcl; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DOMAIN 1 97 CHI.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 222 CH2.
FT DOMAIN 223 329 CH3.
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 113 113 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 143 203
FT DISULFID 249 307
SQ SEQUENCE 329 AA; 36571 MW; 5FCD7B7933850773 CRC64;

Query Match 43.8%; Score 812; DB 1; Length 329;
Best Local Similarity 56.5%; Pred. No. 1.8e-54;
Matches 152; Conservative 49; Mismatches 64; Indels 4; Gaps 2;

Qy 83 LDTVSGSLTYNLTSSDEYEMESPNTDTMKFFLYVDKTHTCPP---CPAPELGGP 138
Db 61 LYTLSSTVTVPSSTWSSQVTCVAHPATKSNLIKRIEPRPKPRPTDICSODNLGRP 120

Qy 139 SVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEVKENYVDGVEVHNAKTPREEQVNS 198
Db 121 SVFIFFPKPKDIIUMITLTPKVTCTVVDVSEEDPVQFSWFVDVNRVFTAQTPHEEQNG 180

Qy 199 TYRVSVLTVHLQDWLNGKEYCKVSKNALPAIEKTIISKAKQPREQVYTLPPSRDEL 258
Db 181 TFRVSVLTVHLQDWLNGKEYCKVSKNALPAIEKTIISKAKQPREQVYTLPPSRDEL 240

Qy 259 TKNOVSLTCLVKGFYPSDIAVWESNGOPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQ 318
Db 241 SKNVSLTCTMVTSTFYPAISISVEWERNGEQDYKNTLPVLDSSEYFLYSKLSVDVDSWM 300

Qy 319 QGVFSCVMHEALHNHYTKSLSLSPGK 347
Db 301 RGDYTCVVHEALHNHHTQKNLSRSPGK 329

RESULT 14
GCCAA_MOUSE GCCAA_MOUSE STANDARD; PRT; 330 AA.
AC P01863;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ig gamma-2A chain C region, A allele.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81076554; PubMed=6777755;
RA Sikorav J.-L., Auffray C., Rougeon F.;
RT "Structure of the constant and 3' untranslated regions of the murine
Raib/c gamma 2a heavy chain messenger RNA.";
RL Nucleic Acids Res. 8:3143-3153(1980).
CC [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=81198976; PubMed=6262729;
RA Yamawaki-Kataoka Y., Miyata T., Honjo T.;
RT "The complete nucleotide sequence of mouse immunoglobulin gamma 2a gene
and evolution of heavy chain genes: further evidence for intervening
sequence-mediated domain transfer.";
RL Nucleic Acids Res. 9:1365-1381(1981).
RN [3]

```

```
RP SEQUENCE FROM N.A.
RX MEDLINE=81223894; PubMed=6787604;
RA Ollio R., Aufrey C., Morchamps C., Rougeon F.;
RT "Comparison of mouse immunoglobulin gamma 2a and gamma 2b chain genes
RT suggests that exons can be exchanged between genes in a multigenic
RT family.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:2442-2446(1981).
RN [4]
RP MYELOMA PROTEIN MOPC 173.
RX MEDLINE=74175517; PubMed=4831970;
RA Bourgois A., Fougereau M., Rocca-Serra J.;
RT "Determination of the primary structure of a mouse IgG2a
RT immunoglobulin:amino-acid sequence of the Fc fragment. Implications
RT for the evolution of immunoglobulin structure and function.";
RL Eur. J. Biochem. 43:423-435(1974).
RN [5]
RP DISULFIDE BONDS.
RX MEDLINE=73056887; PubMed=4555406;
RA de Preval C., Fougereau M.;
RT "Determination of the primary structure of a mouse gamma G2a
RT immunoglobulin. Identification of the disulfide bridges.";
RL Eur. J. Biochem. 30:452-462(1972).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00798; CAA24178.1; .
DR PIR; A02152; G2MSA.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF000047; Ig; 2.
DR SMART; SM00410; Ig_Like; 1.
DR SMART; SM00407; IGc1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
DR Immunoglobulin domain; Immunoglobulin C region.
DR NON_TER 1 15
DR DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
DR DISULFID 27 82 INTERCHAIN (WITH A HEAVY CHAIN).
DR DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
DR DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
DR DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
DR DISULFID 144 204
DR DISULFID 250 308
DR MOD_RES 330 330 REMOVED POST-TRANSLATIONALLY.
DR SEQUENCE 330 AA; 36389 MW; B84361C5445A6864 CRC64;
Query Match 43.7%; Score 811.5; DB 1; Length 330;
Best Local Similarity 54.8%; Pred. No. 1.9e-54;
Matches 155; Conservative 37; Mismatches 60; Indels 31; Gaps 3;
QY 83 LFTVSGSLTIYNLT-----SSDEDEYEMESPNTIDTMKFFLYDKHTHC 126
Db 61 LVTLSSTVTSTWSPQSITCNVAHPASSTKVVKIEPRGPTI-----KPC 107
QY 127 PP-CPAPELGLGSPVLEPPPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV 184
Db 108 PPKRCAPNLGGSPVFIFFPKIKDLVLSLSPITCVVDVSEDPPQVLSWFWNNVEV 167
QY 185 HNAKTKPREQYNTYRVYSLTVLHODWLNKGEYCKVSNKALPAPIETKTISKAKGQPR 244
Db 168 HFAQTQTHREDYNTSLRVYSALPIQHDWMSGKEFKCKVNNKDLPAPIERTISKPKGSVR 227
QY 245 EPQVYVTPPSRDLTKNOYSLCLVKGFPSPDIKAVENSGNCPENNYKTPPEVLDSGSF 304
Db 228 APOVYVLPPEEMTKKQVTLTCWTDWFDPEDIYEWNNKGTENKLTNPVLDSDGSY 287
```

```
QY 305 FLYSKLVTDKSRMOQGNVFCSCVMHEALHNHYTKSLSPGK 347
Db 288 FMYSLKRVKKNWYERNYSVCSVVEGLHNHHTTKFSRTPGK 330
RESULT 15
GCAM_MOUSE
ID GCAM_MOUSE STANDARD; PRT; 399 AA.
AC P01865;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2A chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=82222190; PubMed=6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
RT immunoglobulin gamma chains.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
CC -!- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA
CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
CC BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
CC IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
CC SEGMENT OF MU CHAINS.
CC -!- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 1-329 IS ASSUMED TO BE
CC IDENTICAL WITH THE CORRESPONDING REGION OF THE SECRETED FORM OF
CC THE A ALLELE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00471; AAB59661.1; ALT_INIT.
DR PIR; A02154; G2MSAM.
DR HSSP; P01842; 7FAB.
DR MGD; MGI:96443; Igh-1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF000047; Ig; 2.
DR SMART; SM00410; Ig_Like; 1.
DR SMART; SM00407; IGc1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
DR Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
DR Transmembrane; Alternative splicing.
DR NON_TER 1 15
DR DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
DR DISULFID 27 82 INTERCHAIN (WITH A HEAVY CHAIN).
DR DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
DR DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
DR DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
DR DISULFID 144 204
DR DISULFID 250 308
DR TRANSMEM 346 363
DR DOMAIN 364 399
DR CARBOHYD 180 180
DR SEQUENCE 399 AA; 44020 MW; 4C38138BFAED3FF0 CRC64;
Query Match 43.5%; Score 806.5; DB 1; Length 399;
Best Local Similarity 54.6%; Pred. No. 5.8e-54;
Matches 154; Conservative 37; Mismatches 60; Indels 31; Gaps 3;
QY 83 LDTVSGSLTIYNLT-----SSDEDEYEMESPNTIDTMKFFLYDKHTHC 126
```

```
Db 61 LYTSSSVTVSTWPSQSITCNVAHPASSTKVDKKIEPRGPTI-----KPC 107
QY 127 PP--CPAPELLGSPSVFLFPKPKDTLMISRTEVTCTVVVDVSHEDPEVKFNWYVDGVEV 184
Db 108 PPCKCPAPNLGGPSVFIFFPKIRDVLMLISLSPIVTCVVVDVSEDDPDVQISWFNVEV 167
QY 185 HNAKTPREEOYNSTYRVVSVLTVLHODWLNKKEYCKKVS NKALPAPIEKTISKAKGQPR 244
Db 168 HTAQOTQTHREDYNSTLRVVSALPIQHODWNSGKEFKCKVNNKDLPAPIERTISKPKGSVR 227
QY 245 EPQVYTLPPSRDELTKNOVSLTCLVKGYFSPDIAVEWESNGQPENNYKTTTPPVLDSDGSF 304
Db 228 APQVYVLPPEEEMTKQVTLTCMTDFMPEDIYVEWTNNGKTELNYKNTPEVLDSDGSY 287
QY 305 FLYSKLTVDKSRWQGNVFCGVMHEALHNHYTKRSLSPG 346
Db 288 FMYSKLRVEKKWVERNSYSCSVVHEGLHNHHTTKSFSTPG 329
```

Search completed: January 28, 2003, 08:39:38
Job time : 24.6683 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 28, 2003, 08:38:45 ; Search time 73.2362 Seconds
(without alignments)
976.271 Million cell updates/sec

Title: US-09-730-465-8

Perfect score: 1856

Sequence: 1 MVAGSDAGRALGVLSVVCLL.....MHEALHNHYTKLSLSLSPGK 347

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 205047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phage.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_virus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1233	66.4	471	4 Q8TC77	Q8tc77 homo sapien
2	1233	66.4	701	4 Q96PQ8	Q96pq8 homo sapien
3	1135	61.2	473	4 Q8TC63	Q8tc63 homo sapien
4	909	49.0	337	6 Q95M34	Q95m34 equus caball
5	831.5	44.8	469	11 Q8R3V9	Q8r3v9 mus musculus
6	830.5	44.7	463	11 Q91LC4	Q91lc4 mus musculus
7	826.5	44.5	437	11 Q91A44	Q91a44 mus musculus
8	811	43.7	473	11 Q9D8L4	Q9d8l4 mus musculus
9	807.5	43.5	468	11 Q99L31	Q99l31 mus musculus
10	807.5	43.5	473	11 Q99L25	Q99l25 mus musculus
11	784.5	42.3	473	11 Q91Z05	Q91z05 mus musculus
12	784.5	42.3	474	11 Q8R3H6	Q8r3h6 mus musculus
13	632.5	34.1	240	4 Q9BRW0	Q9brw0 homo sapien
14	476	25.6	134	4 Q14748	Q14748 homo sapien
15	436	23.5	119	4 Q16393	Q16393 homo sapien
16	358.5	19.3	375	4 Q9BSZ1	Q9bsz1 homo sapien

17	358.5	19.3	597	4 Q9BQB8	Q9bqb8 homo sapien
18	358.5	19.3	597	4 Q96BB9	Q96bb9 homo sapien
19	358.5	19.3	597	4 Q9BU10	Q9bu10 homo sapien
20	348.5	18.8	588	4 Q8WUX4	Q8wux4 homo sapien
21	348.5	18.8	613	4 Q96EY0	Q96ey0 homo sapien
22	348.5	18.8	613	4 Q8WUK1	Q8wuk1 homo sapien
23	348.5	18.8	614	4 Q96GAG	Q96gag homo sapien
24	348.5	18.8	618	4 Q96AA6	Q96aa6 homo sapien
25	347	18.7	613	11 Q8VCX7	Q8vcx7 mus musculus
26	329.5	17.8	227	6 Q28754	Q28754 ovis sp. lf
27	329.5	17.8	253	6 Q28753	Q28753 ovis sp. lf
28	319.5	17.2	159	6 Q28752	Q28752 ovis sp. lf
29	280.5	15.1	684	13 Q90544	Q90544 ginglymosto
30	272	14.7	384	4 Q9UP60	Q9up60 homo sapien
31	272	14.7	494	4 Q96K68	Q96k68 homo sapien
32	272	14.7	496	4 Q96KX8	Q96kx8 homo sapien
33	271.5	14.6	416	4 Q9NPP6	Q9npp6 homo sapien
34	269	14.5	496	4 Q96DK0	Q96dk0 homo sapien
35	267	14.4	497	4 Q8WY24	Q8wy24 homo sapien
36	267	14.4	500	4 Q9BRV0	Q9brv0 homo sapien
37	261	14.1	486	11 Q91Z07	Q91z07 mus musculus
38	261	14.1	487	11 Q99KA4	Q99ka4 mus musculus
39	260	14.0	426	11 Q9DCD9	Q9dcd9 mus musculus
40	253.5	13.7	481	11 Q91WT3	Q91wt3 mus musculus
41	253.5	13.7	481	11 Q91WT1	Q91wt1 mus musculus
42	253.5	13.7	482	11 Q91X92	Q91x92 mus musculus
43	253.5	13.7	484	11 Q8VEA0	Q8vea0 mus musculus
44	253.5	13.7	488	11 Q91WR1	Q91wr1 mus musculus
45	253.5	13.7	489	11 Q8VCX4	Q8vcx4 mus musculus

ALIGNMENTS

RESULT 1

Q8TC77 PRELIMINARY; PRT; 471 AA.
 AC Q8TC77;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical 51.8 kDa protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SPLEEN;
 RA Strausberg R.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC024289; AAH24289.1;
 KW Hypothetical protein.
 SQ SEQUENCE 471 AA; 51791 MW; 388F7F4CF588660E CRC64;

Query Match 66.4%; Score 1233; DB 4; Length 471;
 Best Local Similarity 100.0%; Pred. No. 1e-99;
 Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	121	DKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVYVDSHEDPEVKFNWYD	180
Db	245	DKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVYVDSHEDPEVKFNWYD	304
Qy	181	GVEVINATKPREQYNTSTYRVYVSLTVLHQDLNGKEYCKYKVSNNKALPAPEKTISKAK	240
Db	305	GVEVINATKPREQYNTSTYRVYVSLTVLHQDLNGKEYCKYKVSNNKALPAPEKTISKAK	364
Qy	241	GOPREPQVYTLPPSRDELTKNQVSLTCLVKGYGSPSDIAVENESNGPENNYKTTTPVLDS	300
Db	365	GOPREPQVYTLPPSRDELTKNQVSLTCLVKGYGSPSDIAVENESNGPENNYKTTTPVLDS	424
Qy	301	DGSFFLYSKLVDSRWQQQNVFSCSMHEALHNHYTKLSLSLSPGK	347

Db 425 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPGK 471

RESULT 2

Q96PQ8 PRELIMINARY; PRT; 701 AA.

AC Q96PQ8: (1)

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Factor VII active site mutant immunconjugate.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21477448; PubMed=11593034;

RA Hu Z., Garen A.;

RT "targeting tissue factor on tumor vascular endothelial cells and tumor

RT cells for immunotherapy in mouse models of prostatic cancer.";

RL Proc. Natl. Acad. Sci. U.S.A. 98:12180-12185(2001).

DR EMBL; AF272774; AAK58686.1; -.

DR InterPro; IPR000152; Asx_hydroxyl.

DR InterPro; IPR000742; EGF-like.

DR InterPro; IPR001881; EGF_2.

DR InterPro; IPR001881; EGF_Ca.

DR InterPro; IPR003006; Ig_MHC.

DR InterPro; IPR001254; Ser_protease_Try.

DR InterPro; IPR000294; VitK_dep_GLA.

DR Pfam; PF00008; EGF; 2.

DR Pfam; PF00594; gla; 1.

DR Pfam; PF00047; ig; 2.

DR Pfam; PF00089; trypsin; 1.

DR SMART; SM00181; EGF; 2.

DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.

DR PROSITE; PS00022; EGF_1; UNKNOWN_1.

DR PROSITE; PS01186; EGF_2; UNKNOWN_1.

DR PROSITE; PS01187; EGF_Ca; UNKNOWN_1.

DR PROSITE; PS00011; GLU_CARBOXYLATION; UNKNOWN_1.

DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.

DR PROSITE; PS00134; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.

DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.

KW Hydrolase; Serine protease.

SQ SEQUENCE 701 AA; 77826 MW; 94AC6CEB42CC992F CRC64;

Query Match 66.4%; Score 1233; DB 4; Length 701;

Best Local Similarity 100.0%; Pred. No. 1.7e-99;

Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 DKHTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 180

Db 475 DKHTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 534

QY 181 GVEVHNKTKPREEQYNSTYRVSVLTVLHODWLNCKEYKCKVSNKALPAPIETISKAK 240

Db 535 GVEVHNKTKPREEQYNSTYRVSVLTVLHODWLNCKEYKCKVSNKALPAPIETISKAK 594

QY 241 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 300

Db 595 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 654

QY 301 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPGK 347

Db 655 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPGK 701

RESULT 3

Q8TCG3 PRELIMINARY; PRT; 473 AA.

AC Q8TCG3

DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Hypothetical 52.0 kDa protein.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX TISSUE=KIDNEY;

RA Strausberg R.;

RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC025985; AAH25985.1; -.

KW Hypothetical protein.

SQ SEQUENCE 473 AA; 51986 MW; E29920B09BA369F5 CRC64;

Query Match 61.2%; Score 1135; DB 4; Length 473;

Best Local Similarity 93.7%; Pred. No. 4e-91;

Matches 208; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 126 CPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVH 185

Db 252 CPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVH 311

QY 186 NAKTKPREEQYNSTYRVSVLTVLHODWLNCKEYKCKVSNKALPAPIETISKAKGQPRE 245

Db 312 NAKTKPREEQYNSTYRVSVLTVLHODWLNCKEYKCKVSNKALPAPIETISKAKGQPRE 371

QY 246 PQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLDSDGSFF 305

Db 372 PQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLDSDGSFF 431

QY 306 LYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPGK 347

Db 432 LYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPGK 473

RESULT 4

Q95M34 PRELIMINARY; PRT; 337 AA.

AC Q95M34;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Immunoglobulin gamma 1 heavy chain constant region

DE (Fragment).

GN IGHG1

OS Equus caballus (Horse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

OX NCBI_TaxID=9796;

RN [1]

RP SEQUENCE FROM N.A.

RX Wagner B.;

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=98383416; PubMed=9717671;

RA Wagner B., Overesch G., Sheoran A., Holmes M., Richards C.,

RA Leibold W., Radbruch A.;

RT "Organization of the equine immunoglobulin heavy chain constant region

RT genes. III. Alignment of c-mu, c-gamma, c-epsilon and c-alpha genes.";

RL Immunobiology 199:105-119(1998).

DR EMBL; AJ300675; CAC44624.1; -.

DR InterPro; IPR003006; Ig_MHC.

DR Pfam; PF00047; ig; 2.

DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.

FT NON_TER

SQ SEQUENCE 337 AA; 37438 MW; A60BF2B01DEF1F6 CRC64;

Query Match 49.0%; Score 909; DB 6; Length 337;

Best Local Similarity 71.9%; Pred. No. 1.6e-71;

Matches 161; Conservative 32; Mismatches 29; Indels 2; Gaps 1;

```

126 CPPCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 185
Db 114 CPKCPPELLGGPSVFIFFPNPKDTLMITRPEVTCVVVDVSOENPDVKKFNNYMGDEVYR 173
QY 186 NAKTKPREQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIKSKAKGQPRE 245
Db 174 TATTRPREQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIKSKAKGQPRE 233
QY 246 PQVYTLPPSDELTKNQVSLTCLVKGFYPSDIAVEWESNGQP--ENNYKTTTPPVLDSDGS 303
Db 234 PQVYTLAPLAPDELUSKSVSVTCLVKDFYPPENIEWQSNQGPPELETIKYSTTQAQODSDGS 293
QY 304 FFYLSKLTVDKSRWQGNVFGSCVMHEALHNHYTQKSLSLSPCK 347
Db 294 YFLYSLKSLVDRNRWQGGTFTCGVMHEALHNHYTQKNVSKNPGK 337

RESULT 5
Q8R3V9 PRELIMINARY; PRT; 469 AA.
AC Q8R3V9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE Hypothetical 52.0 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024405; AAH24405.1; -.
KW Hypothetical protein.
SQ SEQUENCE 469 AA; 51976 MW; 534793F15D05457 CRC64;

Query Match 44.8%; Score 831.5; DB 11; Length 469;
Best Local Similarity 63.8%; Pred. No. 1.6e-64;
Matches 143; Conservative 41; Mismatches 35; Indels 5; Gaps 2;

QY 126 CPP--CPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVE 183
Db 249 CKPICITVPEV---SSVFIFPPPKPKDTLMITRPEVTCVVVDVSHEDPEVKFNWYVDGVE 305
QY 184 VHNAKTKPREQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIKSKAKGQ 243
Db 306 VHTAQTTPREQNSTFSSVSELPIMHQDWLNGKEYKCKVSNKALPAPIEKTIKSKGRP 365
QY 244 REPOVYTLPPSDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLDSDGS 303
Db 366 KAPOVYIIPPKQMAKDKVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLDSDGS 425
QY 304 FFYLSKLTVDKSRWQGNVFGSCVMHEALHNHYTQKSLSLSPCK 347
Db 426 YFVYSLNVQKSNWEAGTFTCSVLHLEGLHNHHTEKLSLSPCK 469

RESULT 6
Q8R3V9 PRELIMINARY; PRT; 463 AA.
AC Q8R3V9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE Similar to RIKEN cdna 1810060009 gene.
GN IGH-4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
```

```

SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003435; AAH03435.1; -.
DR HSSP; P01842; 7FAB.
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGG1; 2.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_Like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 463 AA; 51007 MW; EAA674C6BBC30783 CRC64;

Query Match 44.7%; Score 830.5; DB 11; Length 463;
Best Local Similarity 56.6%; Pred. No. 1.9e-64;
Matches 151; Conservative 49; Mismatches 62; Indels 5; Gaps 2;

QY 83 LFTVSSGLTIYNLSSDEDEYEMESPNIIDTMKFFLYVDKTHTCPP--CPAPPELLGGPSV 140
Db 200 LITLSSVTVPSSTWTPSETVTCNVHPASSTKVDKKIVPRDCGCKPCICTVPEV---SSV 256
QY 141 FLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREQNSTY 200
Db 257 FIFPPKPKDVLITLTPKVTCTVVVDISKDDPEVQSFNFDVDEVHTAQTPREEQFNSTF 316
QY 201 RVSVSLTVLHODWLNKKEYKCKVSNKALPAPIEKTIKSKAKGQPREPOVYTLPPSDELTK 260
Db 317 RVSELPIMHQDWLNGKEYKCKVSNKALPAPIEKTIKSKAKGQPREPOVYTLPPSDELTK 376
QY 261 NOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLDSDGSFFLYSKLTVDKSRWQ 320
Db 377 DRVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLDSDGSFFLYSKLTVDKSRWQ 436
QY 321 NVFSCSVNHEALHNHYTQKSLSLSPCK 347
Db 437 NTFCTSVLHLEGLHNHHTEKLSLSPCK 463

RESULT 7
Q8R1A4 PRELIMINARY; PRT; 437 AA.
AC Q8R1A4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Gamma1 heavy chain of Mab7 (Fragment).
GN IGH-4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RA Strausberg R.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF152372; AAD40243.1; -.
DR HSSP; P01842; 7FAB.
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_Like; 2.
```

DR	InterPro; IPR003006; Ig_MHC.
DR	InterPro; IPR003596; Ig_V.
DR	Pfam; PF00047; Ig_4.
DR	SMART; SM00409; IG; 2.
DR	SMART; SM00407; IGcl; 3.
DR	SMART; SM00406; IGV; 1.
DR	SMART; SM00410; IG_like; 1.
DR	PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ	SEQUENCE 473 AA; 51699 MW; 9DED57A514475FBB CRC64;

Query Match		43.7%;	Score 811;	DB 11;	Length 473;
Best Local Similarity		64.0%;	Pred. No. 1e-62;		
Matches	142;	Conservative	38;	Mismatches	42;
				Indels	0;
				Gaps	

Qy	126	CPPCAPPELLGGSPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYDGVGEVH	185
Dy	252	CPPCAAPDLLGGSPSFIFFPKTKDVLMSLSLSPMTVCVVVDVEDDPDVQTSFNVNNEVH	311
Qy	186	NAKTPREQYNTSYRVVSVLTVLHDWLNGKEYCKYSKNALPAPIETKTISKAGQPPE	245
Dy	312	TQAOTHTREDYNSTLRVSNALPIQHODMSSKEFKCKVNRRALPSLEKTI SKRPGPVRA	371
Qy	246	POVYTLPSRDELTKNQNSLCVLKGYGPSDIATWENSGOPENNKYTTPVPVLDSDGSFF	305
Dy	372	POQVYLPPLPAEEMTKKEFSLCMTIGFLPAETIAVDMTNGRTEQNKYNTATVLDSDGSYF	431
Qy	306	LXSKLTVDKRWOOGNVFSCSYMHEALHNHYTKSLSLSPGK	347
Dy	432	MYSKLVRQKSTWRGSLFACSVVHEGLHNHLTTKTISRSLCK	473

RESULT 9	
Q99L31	PRELIMINARY; PRT; 468 AA.
ID	Q99L31
AC	Q99L31;
CD	01-JUN-2001 (TREMBLrel. 17, Created)
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE	Similar to RIKEN cDNA 1810060009 gene.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RL	Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
RL	EMBL; BC003878; AAB03878.1; -.
DR	HSSP; P01842; 7FAB.
DR	InterPro; IPR003597; Ig.
DR	InterPro; IPR003597; Ig_c1.
DR	InterPro; IPR003600; Ig_like.
DR	InterPro; IPR003006; Ig_MHC.
DR	InterPro; IPR003596; Ig_V.
DR	Pfam; PF00047; Ig; 3.
DR	SMART; SM00409; IG; 2.
DR	SMART; SM00407; IGcl; 3.
DR	SMART; SM00406; IGV; 1.
DR	SMART; SM00410; IG_like; 1.
DR	PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ	SEQUENCE 468 AA; 51661 MW; 96352328B332ADB CRC64;

Query Match		43.5%;	Score 807.5;	DB 11;	Length 468;
Best Local Similarity		54.8%;	Pred. No. 2e-62;		
Matches	155;	Conservative	37;	Mismatches	60;
				Indels	31;
				Gaps	

Qy	83	LDTVSGSLTYTNUT-----SSDEDEYMESPNTITMKFFLYVDKTHC	126
Dy	199	LXTLSVVVTWSPTWSPQSICNVAHPASSTKVDDKKIPRGFTI-----KPC	245
Qy	127	PP--CPAPELLGGSPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYDGVGEV	185
Dy	246	PCPKCAPNLILGGSPSVIFPFPKTKDVLMSLSLSPMTVCVVVDSEDDPDVQTSFNVNNEVH	305

RESULT 14
Q14748
ID Q14748 PRELIMINARY; PRT: 134 AA.

	Query Match	23.5%	Score 436;	DB 4;	Length 119;
	Best Local Similarity	88.8%;	Prod. No. 1.2e-30;		
	Matches 87;	Conservative	9;	Mismatches	2: Gaps 1:
QY	37	VYGVTFHVPNSVPLKEVLWKKQKOKVAELENSESFRAFSEFKNRYVLTGVSGSLTYINLT	96		
Db	1	VYGVTFHVPNSVPLKEVLWKKQKOKVAELENSESFRAFSEFKNRYVLTGVSGSLTYINLT	60		
QY	97	SSDEDEYEMESPNTDTMTKFFLYV--DKTHTCCPPCAP	132		
Db	61	SSDEDEYEMESPNTDTMTKFFLYVLGHSRHHYALIPTP	98		

Search completed: January 28, 2003, 08:44:18
Job time : 75.2362 secs

10